

EMPFINDLICHKEITSOPTIMIERUNG UND TOLERANZANALYSE ELEKTRISCHER NETZWERKE

von

Dieter Suiter

Erlangen 1977

EMPFINDLICHKEITSOPTIMIERUNG UND TOLERANZANALYSE ELEKTRISCHER NETZWERKE

Dem Fachbereich Ingenieurwissenschaften
der Universität Erlangen-Nürnberg

zur Erlangung des Grades

Doktor der Ingenieurwissenschaften

vorgelegt von

Dieter Suiter

Erlangen 1977

Als Dissertation genehmigt vom
Fachbereich Ingenieurwissenschaften
der Universität Erlangen-Nürnberg

Tag der Einreichung: 19.11.76

Tag der Promotion: 17.02.77

Dekan : Prof. Dr. rer.nat. H.J. Schneider
Berichterstatter : Prof. Dr.-Ing. R. Unbehauen
Prof. Dr. phil.nat. H. Schlitt

INHALTSVERZEICHNIS

	Seite
Einführung	1
I. MULTIPARAMETER-EMPFINDLICHKEIT ERSTER UND ZWEITER ORDNUNG UND IHRE BERECHNUNG	4
1. Allgemeines - Definitionen und Anwendungen der Empfindlichkeit	4
1.1 Empfindlichkeitsarten	4
1.2 Multiparameter-Empfindlichkeit	8
1.3 Empfindlichkeiten zweiter Ordnung	9
1.4 Anwendungsgebiete der Empfindlichkeiten	10
1.5 Berechnung der Empfindlichkeiten	11
2. Berechnung der Empfindlichkeiten aus dem Produkt von Übertragungsfunktionen	14
2.1 Empfindlichkeiten erster Ordnung	14
2.1.1 Empfindlichkeiten bezüglich der Impe- danzen	14
2.1.2 Empfindlichkeit bezüglich der Admittanz	19
2.1.3 Empfindlichkeit bezüglich der Steuer- faktoren gesteuerter Quellen	21
2.2 Empfindlichkeiten zweiter Ordnung	27
2.3 Die praktische Berechnung der Empfindlichk.	28
3. Berechnung der Empfindlichkeiten durch differen- zieren einer in symbolischer Form gegebenen Netz- werkfunktion	29
3.1 Darstellung einer Netzwerkfunktion in symbo- lischer Form	29
3.2 Empfindlichkeiten erster Ordnung in symbo- lischer Form	32
3.3 Empfindlichkeiten zweiter Ordnung in symbo- lischer Form	35

II. DIE EMPFINDLICHKEITSOPTIMIERUNG LINEARER NETZWERKE	38
1. Einführung	38
2. Das Empfindlichkeitsmaß	39
2.1 Vergleich der Empfindlichkeitsmaße	39
2.2 Ein integrales Empfindlichkeitsmaß	42
2.3 Empfindlichkeitsmaß und Netzwerkfunktion	43
3. Die Optimierung	45
3.1 Formulierung des Optimierungsproblems	45
3.2 Die modifizierte Zielfunktion	46
3.3 Der Gewichtungsfaktor R_k	49
3.4 Lösung des unbeschränkten Optimierungs- problems	51
3.4.1 Lösung mit Hilfe eines allgemeinen Optimierungsverfahrens	53
3.4.2 Lösung mit Hilfe eines Optimierungs- verfahrens für Quadratsummen	54
4. Die Berechnung der Empfindlichkeiten und deren Ableitungen	56
4.1 Berechnung der Empfindlichkeiten	56
4.2 Berechnung der Ableitungen der Empfindlichk.	56
4.3 Bemerkung	59
5. Die Freiheitsgrade bei der Empfindlichkeitsopti- mierung	60
6. Praktische Beispiele	67

	Seite
III. TOLERANZANALYSE	73
1. Einführung	73
2. Worst-Case-Verfahren	75
2.1 Linearisierung der Netzwerkfunktion	75
2.2 Die Berechnung des Worst-Case mit Hilfe nichtlinearer Programmierungsverfahren	76
3. Statistische Toleranzanalyse	78
3.1 Das Verfahren nach dem Fehlerfortpflanzungs- gesetz	79
3.2 Das Monte-Carlo-Verfahren	82
4. Die Näherung der Wahrscheinlichkeitsdichte durch Pearsonsche Kurven	84
4.1 Berechnung der Momente der Netzwerkfunktion	85
4.2 Das Pearsonsche Kurvensystem	91
4.2.1 Die Pearsonschen Kurven vom Typ I	97
4.2.2 Die Pearsonschen Kurven vom Typ IV	102
4.2.3 Die Pearsonschen Kurven vom Typ VI	104
4.2.4 Die Pearsonschen Kurven vom Typ II	110
4.2.5 Die Pearsonschen Kurven vom Typ VII	111
4.2.6 Die Pearsonschen Kurven vom Typ III	113
4.2.7 Die Pearsonschen Kurven vom Typ V	117
4.2.8 Die Normalverteilung	120
4.3 Berechnung der statistischen Toleranzgrenzen	123
4.3.1 Algorithmus für die Bestimmung der Toleranzgrenzen	124
5. Anwendungsbeispiele	126
5.1 1. Beispiel	126
5.2 2. Beispiel	133
5.3 3. Beispiel	139

	Seite
ZUSAMMENFASSUNG	148
ANHANG	150
LITERATURVERZEICHNIS	160



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41548420_0015

E I N F Ü H R U N G

Beim Entwurf elektrischer Netzwerke wird zunächst von der Annahme ausgegangen, daß die Bauelemente ideal sind, d.h., daß ihre Sollwerte exakt eingehalten werden. Die theoretisch ermittelte Größe eines Bauelementes läßt sich jedoch nur mit begrenzter Genauigkeit herstellen. Diese fertigungsbedingten Toleranzen zusammen mit den Abweichungen auf Grund der Alterung und der äußeren Einflüsse (z.B. Temperatur, Feuchtigkeit) bewirken, daß die tatsächlichen Netzwerkeigenschaften im allgemeinen von den gewünschten abweichen.

In einem zweiten Entwurfsstadium ist es daher erforderlich, den Einfluß der Bauelementeabweichungen auf die Eigenschaften des Netzwerks zu untersuchen. Man spricht dann von einer "Empfindlichkeitsanalyse" des Netzwerks. Oft genügt es, den Einfluß eines einzelnen Bauelementes auf die Netzwerkeigenschaft zu betrachten, z.B. den Einfluß eines Verstärkungsfaktors auf die Pollage eines aktiven Filters. Im allgemeinen wird jedoch der Einfluß aller Bauelemente auf eine Netzwerkfunktion (z.B. Dämpfung, Phase) in einem vorgegebenen Frequenzbereich zu betrachten sein. Man spricht dann von einer "Multiparameter-Empfindlichkeitsanalyse".

Kann eine bestimmte Vorschrift mit unterschiedlichen Netzwerken realisiert werden, so ist es sinnvoll, diese mit Hilfe eines geeigneten Maßes in Bezug auf ihre "Gesamtempfindlichkeit" miteinander zu vergleichen. Unter anderen möglichen Eigenschaften kann dann auch die Gesamtempfindlichkeit als Kriterium für die Auswahl der günstigsten Schaltung hinzugezogen werden.

Das erste Kapitel dieser Arbeit behandelt die Berechnung der Empfindlichkeiten einer Netzwerkfunktion bezüglich aller Bauelemente. Es werden dabei zwei Verfahren betrachtet, die sich besonders gut für die Anwendungen der Empfindlichkeiten im Rahmen der vorliegenden Arbeit (Empfindlichkeitsoptimierung und Toleranzanalyse) eignen. Dies bedeutet insbesondere, daß mit ihnen auch die Empfindlichkeiten zweiter Ordnung berechnet werden können.

Sind beim Entwurf eines Netzwerks Freiheitsgrade in der Auswahl

der Bauelementsgrößen vorhanden, so können diese dazu genutzt werden, die Gesamtempfindlichkeit zu minimieren. Damit erreicht man, daß die Schaltung gegenüber Schwankungen der Bauelemente-größen unempfindlich wird. Ein Vergleich verschiedener Realisie-rungsmöglichkeiten einer gleichen Vorschrift ist nämlich nur dann sinnvoll, wenn zuvor sichergestellt wurde, daß eine jede Schaltung bereits so dimensioniert wurde, daß sie ihre minimal erreichbare Gesamtempfindlichkeit hat. Das zweite Kapitel be-faßt sich mit dieser Aufgabe der Empfindlichkeitsoptimierung von Netzwerken.

Ist auf Grund eines Gesamtempfindlichkeitsvergleichs und/oder auf Grund anderer Kriterien die Entscheidung für eine bestimmte Schaltung gefallen, so werden in der Regel die maximal zulässig-en Fehler eines jeden Bauelements (die Toleranzen) festgelegt. Hierbei ist darauf zu achten, daß die zulässigen Toleranzgrenzen der Netzwerkeigenschaft nicht überschritten werden. Die Bauelem-mentetoleranzen werden entweder nach dem sog. Schlimmstfall (worst-case) oder mit Hilfe statistischer Verfahren ausgelegt. Meistens wird auch der Gesichtspunkt der Kostenoptimalität mit hinzugezogen. Dieser Schritt im Entwurfsprozeß eines Netzwerkes ist nicht Gegenstand der vorliegenden Arbeit, da er z.B. in [1], [2] und [3] bereits ausführlich behandelt wurde.

Das umgekehrte Problem, die Bestimmung der Toleranzgrenzen einer Netzwerkfunktion bei bekannten Bauelementetoleranzen, nimmt meist den letzten Platz in der Kette von Entwurfsschritten eines Netz-werkes ein. Dieser Aufgabe, der sog. Toleranzanalyse, widmet sich das dritte Kapitel dieser Arbeit.

In der Toleranzanalyse werden im allgemeinen das Worst-Case-Ver-fahren, das Fehlerfortpflanzungsverfahren oder das Monte-Carlo-Verfahren verwendet. Die ersten beiden Methoden beruhen auf der Linearisierung der Netzwerkfunktion und sind naturgemäß nur dort anwendbar, wo die Linearisierung eine hinreichend genaue Nähe-rung liefert. Dies ist z.B. für die Dämpfung in der Umgebung der Anpassungsstellen von Reaktanzfiltern und im Übergangsbereich von einer bestimmten Vorschrift zu einer grundlegend verschiede-nen Vorschrift bereits nicht mehr der Fall. Das Monte-Carlo-Ver-

fahren beruht dagegen nicht auf einer Näherung, sondern auf der Berechnung einer sehr großen Anzahl von Netzwerken, deren Parameter innerhalb der Toleranzgrenzen zufällige Werte annehmen dürfen. Diese Methode führt zwar zu einem genaueren Ergebnis als die ersten beiden, ist aber sehr viel rechenzeit-aufwendiger.

Im dritten Teil dieser Arbeit wird ein weiteres Toleranzanalyseverfahren behandelt, das einerseits genauer ist als das Worst-Case- oder das Fehlerfortpflanzungsverfahren und andererseits schneller ist als das Monte-Carlo-Verfahren. Es beruht auf der Entwicklung der Netzwerkfunktion in eine Taylorreihe zweiter Ordnung. Die Wahrscheinlichkeitsdichtefunktion, die dabei gewonnen wird, ist nicht mehr, wie im Fall der Fehlerfortpflanzungsmethode, die Gaußsche Normalverteilung. Sie kann vielmehr eine beliebige Form haben, die durch das sog. Pearsonsche Kurvensystem angenähert wird. Aus der somit gewonnenen Verteilung der Netzwerkfunktion lassen sich schließlich die gesuchten Toleranzgrenzen berechnen.

Dieses Verfahren ist immer dann anwendbar, wenn die Taylorreihe zweiter Ordnung hinreichend genau die Netzwerkfunktion annähert und wenn für die Beschreibung der Verteilung die Momente bis zur vierten Ordnung ausreichend sind.

Herrn Prof. Dr.-Ing. R. Unbehauen, auf dessen Anregung diese Arbeit zurückgeht, gilt mein besonderer Dank für die wohlwollende Förderung und eingehende Durchsicht der Arbeit.

Herrn Prof. Dr. phil. nat. H. Schlitt danke ich für die Übernahme des Korreferats.

I. MULTIPARAMETER-EMPFINDLICHKEIT ERSTER UND ZWEITER ORDNUNG UND IHRE BERECHNUNG

1. Allgemeines - Definitionen und Anwendungen der Empfindlichkeiten [4], [5]

Im weitesten Sinne werden Empfindlichkeiten als Ableitungen irgendwelcher Netzeigenschaften nach irgendwelchen Bauelementen des Netzwerkes definiert. Die Empfindlichkeiten sind daher ein Maß für den Einfluß der Änderungen der Bauelemente auf Änderungen der Netzeigenschaften. Aus diesem Grund werden sie auch oft "Einflußempfindlichkeiten" genannt.

1.1 Empfindlichkeitsarten

Die Empfindlichkeiten können je nach Art der Ableitung in absolute, halbrelative und vollrelative Empfindlichkeiten klassifiziert werden. Stellt $y(x)$ die Netzeigenschaft in Abhängigkeit des Parameters x dar, dann ist

$$s_x^y = \frac{dy}{dx} \quad (1.1)$$

die absolute Empfindlichkeit,

$$s_x^y = \frac{\frac{dy}{y}}{\frac{dx}{x}} = \frac{xy}{x} = x s_x^y \quad (1.2)$$

die halbrelative Empfindlichkeit und

$$s_x^y = \frac{\frac{dy}{y}}{\frac{dx}{x}} = \frac{xy}{y} = \frac{x}{y} s_x^y \quad (1.3)$$

die vollrelative Empfindlichkeit der Netzeigenschaft $y(x)$ bezüglich des Netzwerkparameters x . Die meist gebrauchte dieser Empfindlichkeitsarten ist die vollrelative. Sie stellt die relative differentieller Änderung der Netzeigenschaft $y(x)$ in Bezug auf die relative differentieller Änderung des Netzwerkparameters x dar. Ist z.B. $s_x^y = 0.5$, dann bedeutet dies, daß eine Änderung von 1 % im Parameterwert eine Änderung von 0.5 %

in der Netzwerkeigenschaft bewirkt.

Die Empfindlichkeiten können auch nach der Art der Netzwerkeigenschaft auf die sie sich beziehen, wie folgt klassifiziert werden:

a) Empfindlichkeit einer allgemeinen Netzwerkfunktion

Als Netzwerkeigenschaft y wird hier eine Netzwerkfunktion $F(x,p)$ (z.B. Übertragungsfunktion, Eingangsimpedanz) betrachtet. Für den Fall der vollrelativen Empfindlichkeit ist dann

$$S_x^F(p) = \frac{\partial \ln F}{\partial \ln x} \quad . \quad (1.4)$$

(Dies ist der reziproke Wert der ursprünglich von Bode [6] definierten Empfindlichkeit).

Für einen bestimmten Frequenzpunkt $p = j\omega$ besitzt die Empfindlichkeit nach Gl. (1.4) im allgemeinen einen komplexen Wert. Dieser Wert stellt daher die relative differentielle Änderung sowohl im Real- als auch im Imaginärteil der Netzwerkfunktion $F(x,p)$ in Bezug auf die relative differentielle Änderung des Netzwerkparameters x dar. Da komplexe Werte als solche nicht direkt verglichen werden können, sondern nur deren Real- und Imaginärteil bzw. deren Betrag und Phase, ist diese Empfindlichkeitsart ungeeignet, wenn verschiedene Netzwerke in Bezug auf ihre Empfindlichkeit verglichen werden sollen. In den meisten Fällen interessiert jedoch entweder nur der Dämpfungs- oder nur der Phasenverlauf eines Netzwerkes. Es würde daher genügen, entweder nur den Realteil oder nur den Imaginärteil der Empfindlichkeit nach Gl. (1.4) zu berechnen.

b) Empfindlichkeit einer reellwertigen Netzwerkfunktion

Weitaus häufiger als die Empfindlichkeit einer komplexwertigen Netzwerkfunktion wird die Empfindlichkeit einer davon abgeleiteten reellen Netzwerkfunktion $f(x,\omega)$ interessieren. In der Regel wird nämlich entweder ein vorgeschriebener Verlauf der Dämpfung oder der Gruppenlaufzeit (bzw. des Betrages oder der Phase) zu realisieren sein.

Wird die Netzwerkfunktion F in Abhängigkeit von Betrag $|F|$ und Phase ϕ ausgedrückt

$$F = |F| e^{j\phi},$$

erhält man durch Einsetzen in Gl. (1.3) folgende nützlichen Beziehungen:

$$S_x^F = S_x^{|F|} + j \phi S_x^\phi \quad (1.4)$$

$$S_x^{|F|} = \operatorname{Re} \{ S_x^F \} \quad (1.5)$$

$$S_x^\phi = \frac{1}{\phi} \operatorname{Im} \{ S_x^F \} \quad (1.6)$$

Für die Dämpfung in Neper $a(\omega) = -\ln |H(j\omega)|$ einer Übertragungsfunktion $H(j\omega)$ ist gemäß Gln. (1.2), (1.3) und (1.5)

$$S_x^{a(\omega)} = -\operatorname{Re} \{ S_x^{H(j\omega)} \}. \quad (1.7)$$

Die halbrelative Empfindlichkeit der Dämpfung in Neper ist daher gleich dem negativen Realteil der vollrelativen Empfindlichkeit der Übertragungsfunktion.

c) Polynom-Empfindlichkeit

In einigen Spezialfällen interessiert die Empfindlichkeit des Zähler- oder Nennerpolynoms einer Netzwerkfunktion. Die vollrelative Polynom-Empfindlichkeit eines Polynoms $P(x,p)$ wird wie folgt definiert:

$$S_x^P = \frac{\partial \ln P}{\partial \ln x}. \quad (1.8)$$

Ist $F(x,p)$ eine Netzwerkfunktion, die durch den Quotienten zwischen Zähler- und Nennerpolynom

$$F(x,p) = \frac{Z(x,p)}{N(x,p)}$$

gegeben ist, so erhält man mit Hilfe der Gln. (1.3) und (1.8) folgende Beziehung zwischen der Empfindlichkeit der Netzwerkfunktion und der des Zähler- und Nennerpolynoms:

$$S_x^F = S_x^Z - S_x^N. \quad (1.9)$$

d) Koeffizienten-Empfindlichkeit

Oft interessiert der Einfluß eines Netzwerkparameters auf die Koeffizienten des Zähler- und Nennerpolynoms einer Netzwerkfunktion. Wenn

$$P(x, p) = \sum_{j=0}^n b_j p^j$$

das interessierende Polynom ist, dann ist

$$S_x^{b_j} = \frac{\partial \ln b_j}{\partial \ln x} \quad (1.10)$$

die vollrelative Koeffizientenempfindlichkeit.

e) Pol- und Nullstellenempfindlichkeit

Es sei

$$F(x, p) = K \frac{\prod_{i=1}^m (p - p_{oi}(x))}{\prod_{j=1}^n (p - p_{\infty j}(x))} \quad (1.11)$$

die interessierende Netzwerkfunktion. Als Pol- bzw. Nullstellenempfindlichkeit wird die halbrelative Empfindlichkeit des Pols $p_{\infty j}(x)$ bzw. der Nullstelle $p_{oi}(x)$ bezüglich des Parameters x definiert

$$S_x^{p_{oi}} = \frac{\partial p_{oi}(x)}{\partial x/x} \quad (1.12)$$

bzw.

$$S_x^{p_{\infty j}} = \frac{\partial p_{\infty j}(x)}{\partial x/x} . \quad (1.13)$$

Durch Einsetzen von Gl. (1.11) in Gl. (1.3) erhält man eine Beziehung zwischen der Empfindlichkeit der Netzwerkfunktion $F(x, p)$ und den Pol- und Nullstellenempfindlichkeiten:

$$S_x^F = S_x^K + \sum_{j=1}^n \frac{1}{p - p_{\infty j}} S_x^{p_{\infty j}} - \sum_{i=1}^m \frac{1}{p - p_{oi}} S_x^{p_{oi}} . \quad (1.14)$$

f) Q- und ω_0 -Empfindlichkeit

Speziell bei aktiven Netzwerken zweiten Grades mit Übertragungsfunktionen der Art

$$H(p) = \frac{Z(p)}{p^2 + 2\sigma_0 p + \omega_0^2}$$

(das Zählerpolynom $Z(p)$ ist höchstens vom zweiten Grad) werden häufig die Empfindlichkeiten des Gütefaktors $Q = \omega_0 / 2\sigma_0$,

$$S_x^Q = \frac{\partial Q / Q}{\partial x / x} \quad (1.15)$$

und des Polbetrages ω_0 ,

$$S_x^{\omega_0} = \frac{\partial \omega_0 / \omega_0}{\partial x / x} \quad (1.16)$$

zur Beurteilung des Netzwerkes herangezogen.

Außer den bereits aufgezeichneten Beziehungen zwischen den verschiedenen Empfindlichkeitsarten, gibt es weitere Beziehungen zwischen den Summen der Empfindlichkeiten aller Bauelemente eines gleichen Netzwerkes. Auf diese sog. Invarianzeigenschaften der Summenempfindlichkeiten, soll in dieser Arbeit nicht eingegangen werden; näheres findet sich z.B. in [3], [4] und [5].

1.2 Multiparameter-Empfindlichkeit

Bisher wurde die Empfindlichkeit einer Netzwerkeigenschaft $y(x)$ nur bezüglich eines einzelnen Parameters x des Netzwerkes betrachtet. Für kleine Änderungen Δx_i aller Netzwerkparameter x_i ($i = 1, 2, \dots, n$), kann die Änderung der Netzwerkeigenschaft Δy durch die linearisierte Taylorentwicklung

$$\Delta y = \sum_{i=1}^n \frac{\partial y}{\partial x_i} \Delta x_i \quad (1.17)$$

approximativ dargestellt werden. Wird $\partial y / \partial x_i$ nach Gl. (1.3) in Gl. (1.17) eingesetzt, ist

$$\frac{\Delta Y}{Y} = \sum_{i=1}^n S_{x_i}^Y \frac{\Delta x_i}{x_i} \quad (1.18)$$

Die relative Änderung der Netzwerkeigenschaft y kann gemäß Gl. (1.18) als eine mit der vollrelativen Empfindlichkeit gewichteten Summe der relativen Änderungen der Netzwerkparameter x_i betrachtet werden.

Werden die Empfindlichkeiten $S_{x_i}^Y = \frac{\partial Y/Y}{\partial x_i/x_i}$ ($i = 1, 2, \dots, n$) in Gl. (1.18) zum Vektor

$$\underline{S}^Y = (S_{x_1}^Y, S_{x_2}^Y, \dots, S_{x_n}^Y) \quad (1.19)$$

und die relativen Änderungen der Bauelemente zum Vektor

$$\underline{\Delta x} = \left(\frac{\Delta x_1}{x_1}, \frac{\Delta x_2}{x_2}, \dots, \frac{\Delta x_n}{x_n} \right) \quad (1.20)$$

zusammengefaßt, dann kann Gl. (1.18) auch als Skalarprodukt dieser beiden Vektoren interpretiert werden

$$\frac{\Delta Y}{Y} = \underline{S}^Y \cdot \underline{\Delta x} \quad (1.21)$$

Auf Grund der Analogie zwischen Gl. (1.21) und Gl. (1.17) für den Fall, daß die Netzwerkeigenschaft nur von einem Parameter abhängt, wird der Vektor $\underline{S}^Y$, Gl. (1.19), von Goldstein und Kuo [7] als Multiparameter-Empfindlichkeit bezeichnet. In diesem Fall handelt es sich um die vollrelative Multiparameter-Empfindlichkeit. In ähnlicher Weise können selbstverständlich auch absolute und halbrerelative Multiparameter-Empfindlichkeiten definiert werden.

1.3 Empfindlichkeiten zweiter Ordnung

Die Terme höherer Ordnung in der Taylor-Entwicklung der Netzwerkfunktion können nicht immer gegenüber den Termen erster Ordnung vernachlässigt werden. In solchen Fällen liefern die Gln. (1.17) und (1.18) ungenügende Näherungen. Es ist dann zweckmäßig, die Terme zweiter Ordnung hinzuzuziehen. Aus Gl. (1.17) wird damit

$$\Delta y = \sum_{i=1}^n \frac{\partial y}{\partial x_i} \Delta x_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 y}{\partial x_i \partial x_j} \Delta x_i \Delta x_j . \quad (1.22)$$

Die zweite Ableitung der Netzwerkeigenschaft $\partial^2 y / \partial x_i \partial x_j$ wird absolute Empfindlichkeit zweiter Ordnung genannt

$$S_{x_i x_j}^y = \frac{\partial^2 y}{\partial x_i \partial x_j} . \quad (1.23)$$

Analog zu der Empfindlichkeit erster Ordnung gibt es auch hier die halbrelative Empfindlichkeit zweiter Ordnung

$$S'_{x_i x_j}^y = x_i x_j \frac{\partial^2 y}{\partial x_i \partial x_j} \quad (1.24a)$$

und die vollrelative Empfindlichkeit zweiter Ordnung

$$S''_{x_i x_j}^y = \frac{x_i x_j}{y} \frac{\partial^2 y}{\partial x_i \partial x_j} . \quad (1.24b)$$

1.4 Anwendungsgebiete der Empfindlichkeiten

Als klassisches Anwendungsgebiet der Empfindlichkeiten steht an erster Stelle die Beurteilung der Güte eines Netzwerkes und der Vergleich mit anderen, die die gleiche Netzwerkeigenschaft realisieren.

Den einfachsten Fall stellen dabei die aktiven Netzwerke zweiten Grades dar. Hier werden in der Regel die ω_0 - und Q-Empfindlichkeiten als Gütemaß herangezogen. Oft wird dabei fälschlicherweise nur die Empfindlichkeit bezüglich des aktiven Bauelementes betrachtet, obwohl die Empfindlichkeiten bezüglich der passiven Elemente durchaus die größeren sein können. Im allgemeinen Fall werden jedoch die Empfindlichkeiten einer Netzwerkfunktion (z.B. Dämpfung, Phase) in einem Frequenzintervall ($\omega_1 \omega_2$) gegenüber sämtlichen Bauelementen interessieren. Für Vergleichszwecke wird dann aus den Einzelempfindlichkeiten ein sog. Gesamtempfindlichkeitsmaß berechnet. Auf die Gesamtempfindlichkeitsmaße wird in Kap. II noch näher eingegangen.

Ein weiteres Anwendungsgebiet der Empfindlichkeiten ist die Empfindlichkeitsoptimierung eines Netzwerkes. Hierzu werden außer den Empfindlichkeiten erster Ordnung auch diejenigen zweiter Ordnung benötigt. Wie im Kap. II gezeigt wird, geht es dabei um die Minimierung einer der o.g. Gesamttempfindlichkeitsmaße.

Auch bei anderen Optimierungsverfahren, wie z.B. in der Verlustoptimierung, der iterativen Synthese, der kostenoptimalen Zuweisung der Bauelementetoleranzen werden die Empfindlichkeiten benötigt.

Die Toleranzanalyse, d.h. die Berechnung von Fehlergrenzen der Netzwerkeigenschaft wenn die Bauelemente innerhalb ihrer Toleranzgrenzen schwanken, ist ebenfalls ein sehr weites und häufig benutztes Anwendungsgebiet der Empfindlichkeiten. Kap. III beschäftigt sich mit dieser Aufgabe.

1.5 Berechnung der Empfindlichkeiten

Im Rahmen dieser Arbeit werden primär solche Netzwerke behandelt, die eine Vorschrift im Frequenzbereich, wie z.B. einen Dämpfungs- oder Phasenverlauf zu erfüllen haben. Von den verschiedenen oben aufgeführten Empfindlichkeiten werden daher von nun an nur noch die ersten beiden Arten interessieren, d.h. die Empfindlichkeiten einer allgemeinen und die einer reellwertigen Netzwerkfunktion. Die Empfindlichkeiten einer reellen Netzwerkfunktion $f(\omega)$ lassen sich mit Hilfe der Gln. (1.4) bis (1.7) aus denen einer allgemeinen, $F(p)$ berechnen. Außerdem werden die Empfindlichkeiten bezüglich mehrerer oder sogar aller Parameter eines Netzwerkes, und zwar nicht nur an einem Frequenzpunkt sondern möglichst als Funktion der Frequenz ω zu berechnen sein. Es wird also die Multiparameter-Empfindlichkeit aus Gl. (1.19) benötigt, in der jedes Vektorelement eine Funktion von ω ist

$$\underline{S}^f(\omega) = (S^f_{x_1}(\omega), S^f_{x_2}(\omega), \dots, S^f_{x_n}(\omega)) \quad (1.25)$$

In der Praxis wird der Frequenzbereich $\{\omega_1, \omega_2\}$ mit diskreten Fre-

quenzpunkten ω_v ($v = 1, 2, \dots, m$) abgetastet. Es würde daher genügen, an Stelle der Multiparameter-Empfindlichkeitsfunktionen die Multiparameter-Empfindlichkeiten für jeden Frequenzpunkt zu berechnen. Wie sich bei den verschiedenen Anwendungsfällen der Empfindlichkeiten im Verlauf dieser Arbeit zeigen wird, ist es jedoch sinnvoller, erst die Empfindlichkeiten als Funktionen der Frequenz zu berechnen, um sie dann an den interessierenden Frequenzpunkten auszuwerten. Dies hat den Vorteil, daß der Prozeß der Empfindlichkeitsberechnung von der Anwendung (z.B. Berechnung der Gesamtempfindlichkeit oder Toleranzanalyse) getrennt werden kann. Bei der numerischen Berechnung auf digitalen Rechenanlagen bedeutet dies eine Einsparung an Speicherkapazität in der Zentraleinheit und meistens auch eine Einsparung an Rechenzeit, da in der Regel das Auswerten einer Empfindlichkeitsfunktion schneller als die numerische Analyse eines Netzwerkes zwecks Berechnung der Empfindlichkeit sein wird. Hier werden also folgende Forderungen an ein Verfahren zur Berechnung der Empfindlichkeiten gestellt:

- (a) Berechnung der Empfindlichkeiten einer im allgemeinen komplexen Netzwerkfunktion.
- (b) Berechnung der Empfindlichkeiten gegenüber allen Netzwerkparametern gleichzeitig, d.h. mit möglichst wenig Analyseschritten.
- (c) Berechnung der Empfindlichkeiten als Funktionen des komplexen Frequenzparameters p oder der Kreisfrequenz ω .

Aus der Literatur sind mehrere Verfahren für die Empfindlichkeitsberechnung bekannt. Eine Übersicht darüber sowie weitere Literaturangaben findet man z.B. in [5], [8] und [9]. Einige dieser Verfahren eignen sich nur für die Berechnung der Empfindlichkeit von jeweils einem Bauelement. Ein solches Verfahren müßte also für jedes Bauelement wiederholt werden. Andere Verfahren berechnen die Empfindlichkeiten rein numerisch, d.h. für einzelne Frequenzpunkte. Die meisten dieser Verfahren benötigen jedoch zwei Netzwerkanalysen für die Empfindlichkeitsberechnung, so zum Beispiel auch die häufig verwendete Methode des "adjoint network"

von Director & Rohrer [10]. Im Abschnitt 2 dieses Kapitels wird ein Verfahren behandelt, welches nicht die genannten Nachteile besitzt, alle o.g. Forderungen erfüllt und außerdem auch für die Berechnung der Empfindlichkeiten zweiter Ordnung geeignet ist.

Bei der Optimierung der Gesamtempfindlichkeit oder bei anderen Optimierungsaufgaben, die mit einem Gradientenverfahren arbeiten, müssen die Empfindlichkeiten erster Ordnung (häufig auch die der zweiten Ordnung) nicht nur für unterschiedliche Frequenzpunkte, sondern auch für unterschiedliche Werte der Netzwerkparameter berechnet werden. Außer den oben aufgeführten Forderungen (a) bis (c) muß dann noch die Bedingung

(d) Berechnung der Empfindlichkeiten als Funktion der Bauelemente

erfüllt werden.

Diese letzte Forderung läßt sich nur, wie in Abschnitt 3 dieses Kapitels gezeigt wird, mit der direkten Ableitung bezüglich der Bauelemente einer Netzwerkfunktion in symbolischer Form erfüllen.

2. Berechnung der Empfindlichkeiten aus dem Produkt von Übertragungsfunktionen

Die Empfindlichkeit einer Netzwerkfunktion bezüglich eines Bauelementes kann aus dem Produkt von zwei speziellen Übertragungsfunktionen des Netzwerks berechnet werden. Diese Möglichkeit wurde bereits von Bode [6] für einige Spezialfälle angewandt und später auch von anderen Autoren auf verschiedenste Weise abgeleitet und erweitert [11], [12], [13].

Hier soll nun anhand des Kompensationstheorems gezeigt werden, daß nach o.g. Prinzip und bei Anwendung eines speziellen Netzwerkanalyseprogrammes die Empfindlichkeiten erster und zweiter Ordnung bezüglich aller Netzwerkparameter als Funktionen von p in einem einzigen Analyseschritt gewonnen werden können. Es soll sich dabei nur um lineare, zeitinvariante Netzwerke handeln, die aus Widerständen, Induktivitäten, Kapazitäten und der vier Arten von gesteuerten Quellen aufgebaut sind.

In den folgenden Ausführungen soll unter dem Begriff "Empfindlichkeit" immer die absolute Empfindlichkeit erster Ordnung verstanden werden, es sei denn, Art und Ordnung werden ausdrücklich angegeben.

2.1 Empfindlichkeiten erster Ordnung

2.1.1 Empfindlichkeiten bezüglich der Impedanzen

Man betrachte zunächst das Netzwerk N aus Bild 1.1. Es sind die Empfindlichkeiten einer Netzwerkfunktion, beispielsweise der Leerlauf-Übertragungsfunktion

$$H = \left. \frac{U_b}{U_a} \right|_{I_b = 0} \quad (1.26)$$

bezüglich aller n passiven Elemente zu berechnen. Diese Elemente wurden aus dem mit $N^{(1)}$ bezeichneten Teilnetzwerk herausgezogen

und durch ihre Impedanzen Z_i dargestellt. In $N^{(1)}$ verbleiben daher nur noch die eventuell vorhandenen gesteuerten Quellen.

Es wird angenommen, daß alle n Elemente um ein ΔZ_i verändert werden, wobei $|\Delta Z_i| \ll |Z_i|$ für $i = 1, 2, \dots, n$ sein soll. Dadurch ändern sich sämtliche Spannungen und Ströme im Netzwerk. Die geänderten Werte werden durch einen Strich (') gekennzeichnet. Mit Hilfe des Kompensationstheorems [14] lassen sich die Änderungen der Impedanzen durch Hilfsquellen mit den Werten

$$U_i = \Delta Z_i I'_i, \quad \text{für } i = 1, 2, \dots, n \quad (1.27)$$

gemäß Bild 1.2 darstellen.

Es soll nun die Änderung ΔH in der Übertragungsfunktion in Abhängigkeit von der Änderung der Impedanzen berechnet werden:

$$\Delta H = H' - H \quad \text{mit} \quad H' = \frac{U'_b}{U'_a} \quad \text{bei} \quad I'_b = 0.$$

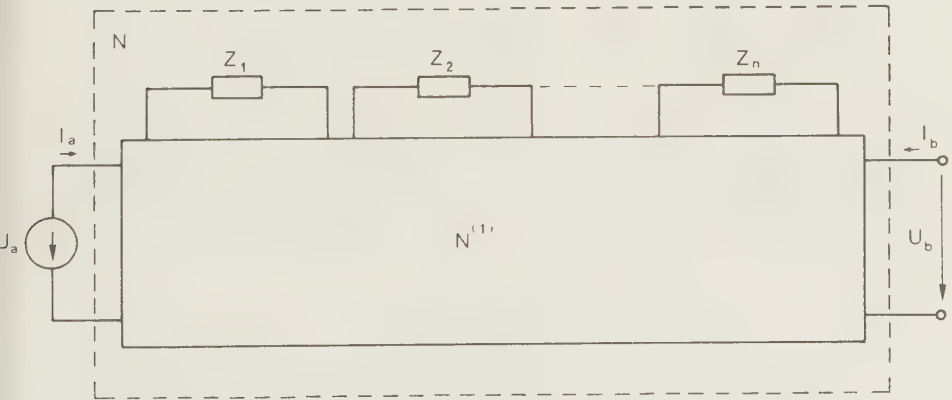


Bild 1.1: Zur Berechnung der Empfindlichkeiten bezüglich der Impedanzen

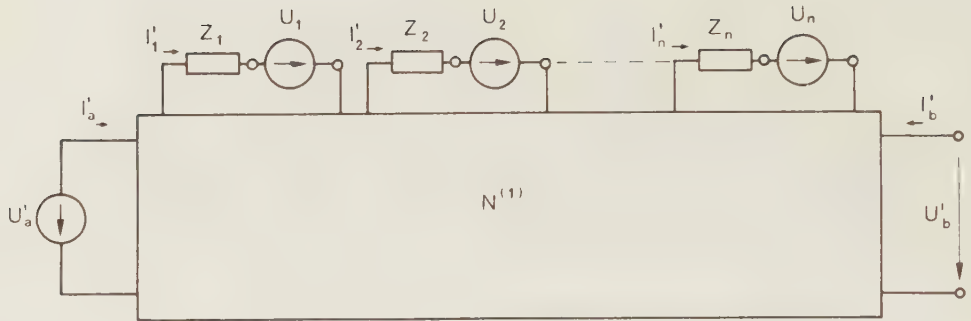


Bild 1.2: Ersetzen der Impedanzänderungen durch Hilfsquellen

Da es sich hier um ein lineares Netzwerk handelt, können auf Grund des Überlagerungssatzes folgende Beziehungen zwischen Ursachen und Wirkungen aufgestellt werden:

$$\begin{bmatrix} U'_b \\ I'_1 \\ I'_2 \\ \vdots \\ I'_n \end{bmatrix} = \begin{bmatrix} h_{ba} & h_{b1} & h_{b2} & \dots & h_{bn} \\ h_{1a} & h_{11} & h_{12} & & h_{1n} \\ h_{2a} & h_{21} & & \dots & \\ \vdots & & & & \\ h_{na} & h_{n1} & & & h_{nn} \end{bmatrix} \cdot \begin{bmatrix} U'_a \\ U_1 \\ U_2 \\ \vdots \\ U_n \end{bmatrix} \quad (1.28)$$

oder mit der hier angegebenen Matrixunterteilung

$$\begin{bmatrix} U'_b \\ \underline{I}' \end{bmatrix} = \begin{bmatrix} \underline{H}_{11} & \underline{H}_{1e} \\ \underline{H}_{e1} & \underline{H}_{ee} \end{bmatrix} \cdot \begin{bmatrix} U'_a \\ \underline{U} \end{bmatrix} \quad (1.29)$$

Die Gln. (1.27) für die Darstellung der Hilfsquellen können nun in die Matrixform

$$\underline{U} = \underline{\Delta Z} \underline{I}' \quad (1.30)$$

gebracht werden, wobei

$$\underline{\Delta Z} = \text{diag} \{ \Delta Z_1, \Delta Z_2, \dots, \Delta Z_n \} \quad (1.31)$$

ist. Wird $\underline{I}'$ aus Gl. (1.29) in Gl. (1.30) eingesetzt und diese nach $\underline{U}$ aufgelöst, erhält man, unter der Voraussetzung, daß die Matrix $(\underline{E} - \underline{\Delta Z} \underline{H}_{ee})$ nichtsingulär ist,

$$\underline{U} = (\underline{E} - \underline{\Delta Z} \underline{H}_{ee})^{-1} \underline{\Delta Z} \underline{H}_{e1} \underline{U}'_a .$$

Wird diese Gleichung wiederum in die erste der Gleichungen (1.29) eingesetzt und nach $\underline{U}'_b / \underline{U}'_a$ aufgelöst, folgt

$$\underline{H}' = \frac{\underline{U}'_b}{\underline{U}'_a} = \underline{H}_{11} + \underline{H}_{1e} (\underline{E} - \underline{\Delta Z} \underline{H}_{ee})^{-1} \underline{\Delta Z} \underline{H}_{e1} . \quad (1.32)$$

Die Matrix $(\underline{E} - \underline{\Delta Z} \underline{H}_{ee})^{-1}$ läßt sich bei hinreichend kleinem $|\Delta Z_i|$ ($i = 1, 2, \dots, n$) in die Reihe

$$(\underline{E} - \underline{\Delta Z} \underline{H}_{ee})^{-1} = \underline{E} + \underline{\Delta Z} \underline{H}_{ee} + (\underline{\Delta Z} \underline{H}_{ee})^2 + (\underline{\Delta Z} \underline{H}_{ee})^3 + \dots \quad (1.33)$$

entwickeln. Durch Einsetzen von Gl. (1.33) in Gl. (1.32) ist

$$\underline{H}' = \underline{H}_{11} + \underline{H}_{1e} \underline{\Delta Z} \underline{H}_{e1} + \underline{H}_{1e} \underline{\Delta Z} \underline{H}_{ee} \underline{\Delta Z} \underline{H}_{e1} + \underline{H}_{1e} (\underline{\Delta Z} \underline{H}_{ee})^2 \underline{\Delta Z} \underline{H}_{e1} + \dots \quad (1.34)$$

Unter der Voraussetzung, daß alle $|\Delta Z_i|$ ($i = 1, 2, \dots, n$) hinreichend klein sind, können in der Gl. (1.34) alle Terme zweiter und höherer Ordnung in $\underline{\Delta Z}$ vernachlässigt werden. Damit ist in erster Näherung

$$\underline{H}' = \underline{H}_{11} + \underline{H}_{1e} \underline{\Delta Z} \underline{H}_{e1} . \quad (1.35)$$

Gemäß Gl. (1.28) und (1.29) ist $\underline{H}_{11} = h_{ba} = \frac{\underline{U}'_b}{\underline{U}'_a}$ bei $\underline{U}_1, \underline{U}_2, \dots, \underline{U}_n$, $\underline{I}_b = 0$, d.h. $\underline{H}_{11} = \underline{H}$ ist die ursprüngliche Übertragungsfunktion aus Gl. (1.26). Damit ergibt sich aus Gl. (1.35)

$$\Delta H = \underline{H}' - \underline{H} = \underline{H}_{1e} \underline{\Delta Z} \underline{H}_{e1} . \quad (1.36)$$

Nach Einsetzen der Untermatrizen $\underline{H}_{1e}$ und $\underline{H}_{e1}$ aus Gln. (1.28) und (1.29) und $\underline{\Delta Z}$ aus Gl. (1.31) in Gl. (1.36) ist

$$\Delta H = (h_{b1}, h_{b2}, \dots, h_{bn}) \begin{bmatrix} \Delta Z_1 & & & \\ & \Delta Z_2 & & \\ & & \ddots & \\ & & & \Delta Z_n \end{bmatrix} \begin{bmatrix} h_{1a} \\ h_{2a} \\ \vdots \\ h_{na} \end{bmatrix}$$

$$\Delta H = h_{1a} h_{b1} \Delta Z_1 + h_{2a} h_{b2} \Delta Z_2 + \dots + h_{na} h_{bn} \Delta Z_n \quad (1.37)$$

Gl. (1.37) kann aber auch als Skalarprodukt der beiden Vektoren

$$\underline{S} = (h_{1a} h_{b1}, h_{2a} h_{b2}, \dots, h_{na} h_{bn}) \quad (1.38)$$

und $\underline{\Delta Z} = (\Delta Z_1, \Delta Z_2, \dots, \Delta Z_n)$

interpretiert werden:

$$\Delta H = \underline{S} \cdot \underline{\Delta Z} = \sum_{i=1}^n h_{ia} h_{bi} \Delta Z_i$$

Wird diese Gleichung mit Gl. (1.17) verglichen, erkennt man, daß die Elemente des Vektors $\underline{S}$ die absoluten Empfindlichkeiten erster Ordnung der Übertragungsfunktion H bezüglich der Impedanzen Z_i

$$S_{Z_i}^H = \frac{\partial H}{\partial Z_i} = h_{ia} h_{bi} \quad (i = 1, 2, \dots, n) \quad (1.39)$$

sind. Der Vektor $\underline{S}$, Gl. (1.38), stellt daher die absolute Multiparameter-Empfindlichkeit dar (vgl. Kap. I, Abschnitt 1.2).

In Gl. (1.39) bedeutet, wie aus Gl. (1.28) ersichtlich,

$$h_{ia} = \frac{I_i'}{U_a'} \quad \text{mit} \quad U_j = 0 \quad (j = 1, 2, \dots, n),$$

die Übertragungsfunktion vom Eingang zum i -ten Element und

$$h_{bi} = \frac{U_b'}{U_i'} \quad \text{mit} \quad U_a' = 0 \quad \text{und} \quad U_j = 0 \quad (j=1, 2, \dots, n; j \neq i),$$

die Übertragungsfunktion vom i -ten Element zum Ausgang.

Die Elemente der Multiparameter-Empfindlichkeit, Gl. (1.38), lassen sich also aus den Produkten von jeweils zwei Übertragungs-

funktionen des Netzwerks aus Bild 1.2 berechnen. Diese Übertragungsfunktionen befinden sich in der ersten Zeile bzw. in der ersten Spalte der Hybridmatrix aus Gl. (1.28).

In völlig analoger Weise kann gezeigt werden, daß auch die Empfindlichkeiten beliebiger Netzwerkfunktionen bezüglich der Impedanz oder Admittanz passiver Netzwerkelemente oder bezüglich der Verstärkungsfaktoren von gesteuerten Quellen als Produkt von zwei Übertragungsfunktionen berechnet werden können.

Im folgenden soll für den einfacheren Fall eines einzelnen Bauelementes gezeigt werden, wie die Empfindlichkeit einer Netzwerkfunktion bezüglich der Admittanz oder des Verstärkungsfaktors der vier Arten von gesteuerten Quellen ebenfalls aus dem Produkt von zwei Übertragungsfunktionen berechnet werden kann. Die gleichzeitige Berechnung der Empfindlichkeiten bezüglich aller Bauelemente geschieht dann durch direkte Erweiterung des Einparameter-Falls auf den Fall mehrerer Parameter.

2.1.2 Empfindlichkeit bezüglich der Admittanz

Die Änderung ΔY_i in der Admittanz eines passiven Bauelementes kann mit Hilfe des Kompensationstheorems durch eine Stromquelle mit dem Wert

$$I_i = \Delta Y_i U'_i \quad (1.40)$$

gemäß Bild 1.3 dargestellt werden.

Nach dem Überlagerungssatz gilt

$$\begin{bmatrix} U'_b \\ U'_i \end{bmatrix} = \begin{bmatrix} h_{ba} & h_{bi} \\ h_{ia} & h_{ii} \end{bmatrix} \cdot \begin{bmatrix} U'_a \\ I_i \end{bmatrix}. \quad (1.41)$$

Aus den Gln. (1.40) und (1.41) folgt, unter Voraussetzung, daß $1 - \Delta Y_i h_{ii} \neq 0$ ist, nach einigen Zwischenschritten:

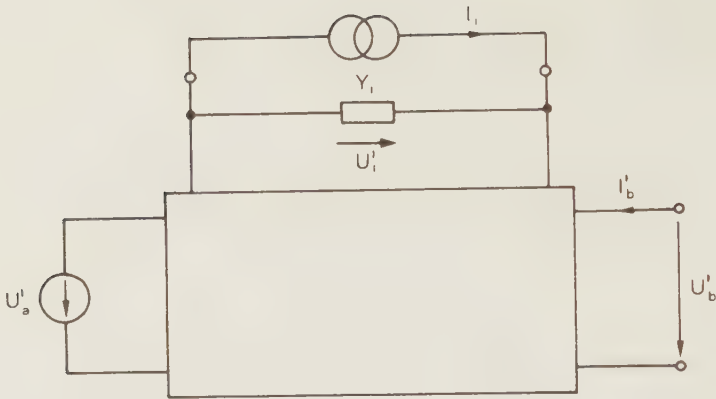


Bild 1.3: Zur Berechnung der Empfindlichkeit bezüglich der Admittanz

$$H' = \frac{U'_b}{U'_a} = h_{ba} + h_{ia} h_{bi} \frac{\Delta Y_i}{1 - \Delta Y_i h_{ii}}$$

und mit $\Delta H = H' - H = H' - h_{ba}$,

$$\frac{\partial H}{\partial Y_i} = \lim_{\Delta Y_i \rightarrow 0} \frac{\Delta H}{\Delta Y_i} = h_{ia} h_{bi}.$$

Daher gilt für die Empfindlichkeit der Übertragungsfunktion bezüglich der Admittanz Y_i

$$S_{Y_i}^H = h_{ia} h_{bi} \quad (1.42)$$

wobei $h_{ia} = \frac{U'_i}{U'_a}$ mit I_i und $I'_b = 0$

die Übertragungsfunktion vom Eingang zum Element ist, und

$$h_{bi} = \frac{U'_b}{I_i} \quad \text{mit} \quad U'_a \text{ und } I'_b = 0$$

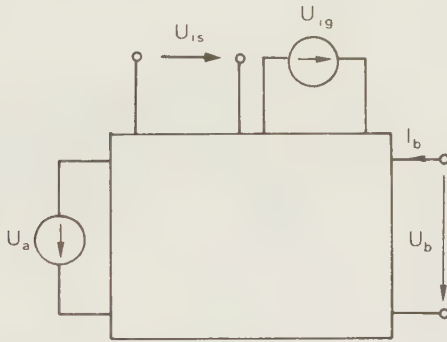
die Übertragungsfunktion vom Element zum Ausgang ist.

2.1.3 Empfindlichkeit bezüglich der Steuerfaktoren gesteuerter Quellen

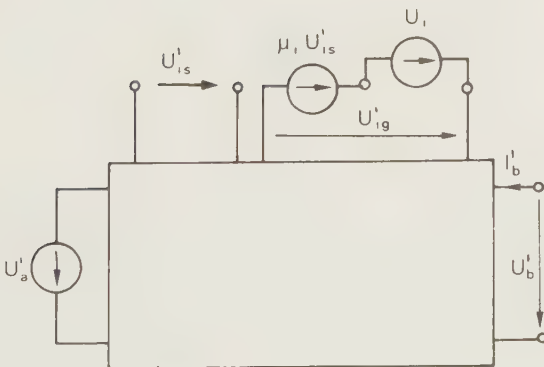
In Bild 1.4a wird eine aus dem Netzwerk herausgezogene spannungs-
gesteuerte Spannungsquelle gezeigt. Die Spannung U_{ig} des gesteu-
erten Zweiges hängt über den Steuerparameter μ_i von der Spannung
des steuernden Zweiges U_{is} nach der Beziehung

$$U_{ig} = \mu_i U_{is} \quad (1.43)$$

ab.



(a)



(b)

Bild 1.4: Zur Berechnung der Empfindlichkeit bezüglich des
Steuerfaktors einer spannungsgesteuerten Spannungs-
quelle

Ändert sich nun der Steuerparameter um den Wert $\Delta\mu_i$, so ändern sich sämtliche Spannungen und Ströme im Netzwerk. Insbesondere wird aus U_{ig}

$$U'_{ig} = \mu_i U'_{is} + \Delta\mu_i U'_{is} . \quad (1.44)$$

Die Änderung des Steuerparameters kann daher, wie in Bild 1.4 (b) gezeigt, durch eine mit der gesteuerten Quelle in Reihe geschalteten Hilfsquelle mit dem Wert

$$U_i = \Delta\mu_i U'_{is} \quad (1.45)$$

dargestellt werden.

Es soll nun die Empfindlichkeit der Leerlaufübertragungsfunktion $H = U_b/U_a$ gegenüber dem Steuerparameter μ_i berechnet werden. Nach dem Überlagerungssatz gilt für Bild 1.4 (b) die Matrixgleichung

$$\begin{bmatrix} U'_b \\ U'_{is} \end{bmatrix} = \begin{bmatrix} h_{ba} & h_{bi} \\ h_{ia} & h_{ii} \end{bmatrix} \begin{bmatrix} U'_a \\ U_i \end{bmatrix} . \quad (1.46)$$

Aus den Gln. (1.45) und (1.46) folgt unter der Voraussetzung, daß $1 - \Delta\mu_i h_{ii} \neq 0$ ist, nach einigen Zwischenschritten

$$H' = h_{ba} + h_{ia} h_{bi} \frac{\Delta\mu_i}{1 - \Delta\mu_i h_{ii}}$$

und mit $\Delta H = H' - H = H' - h_{ba}$,

$$\frac{\partial H}{\partial \mu_i} = \lim_{\Delta\mu_i \rightarrow 0} \frac{\Delta H}{\Delta\mu_i} = h_{ia} h_{bi} .$$

Daher gilt für die Empfindlichkeit der Übertragungsfunktion bezüglich des Steuerparameters μ_i

$$S_{\mu_i}^H = h_{ia} h_{bi} , \quad (1.47)$$

wobei

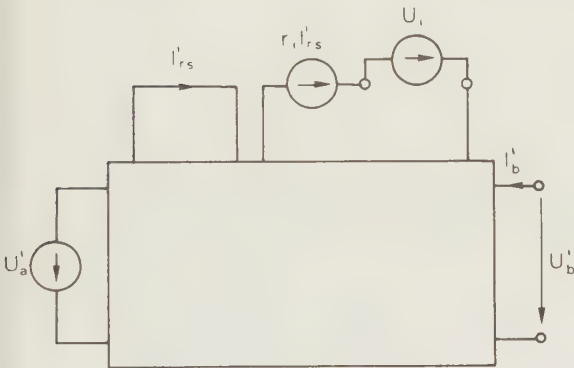
$$h_{ia} = \frac{U'_{is}}{U'_a} \quad \text{mit} \quad U_i \quad \text{und} \quad I'_b = 0$$

die Übertragungsfunktion vom Eingang zum steuernden Zweig und

$$h_{bi} = \frac{U'_b}{U'_i} \quad \text{mit} \quad U'_a \quad \text{und} \quad I'_b = 0$$

die Übertragungsfunktion vom gesteuerten Zweig zum Ausgang ist.

Für die restlichen drei Arten von gesteuerten Quellen wird in ähnlicher Weise vorgegangen. Lage und Art der Hilfsquelle und die notwendigen Übertragungsfunktionen für die Berechnung der Empfindlichkeit in jedem dieser Fälle, können aus Bild 1.5 entnommen werden.

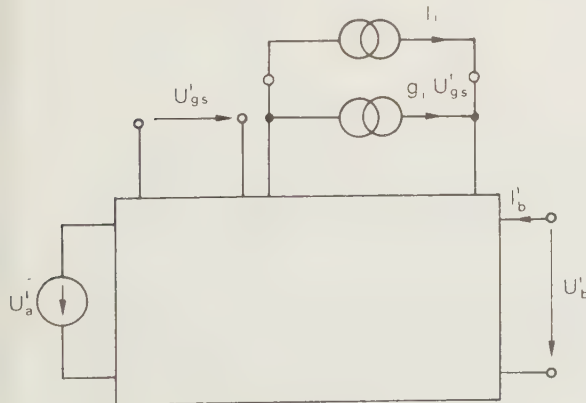


$$S_{r_i}^H = h_{ia} h_{bi}$$

$$h_{ia} = \left. \frac{I'_{rs}}{U'_a} \right|_{U'_i, I'_b = 0}$$

$$h_{bi} = \left. \frac{U'_b}{U'_i} \right|_{U'_a, I'_b = 0}$$

a) Stromgesteuerte Spannungsquelle

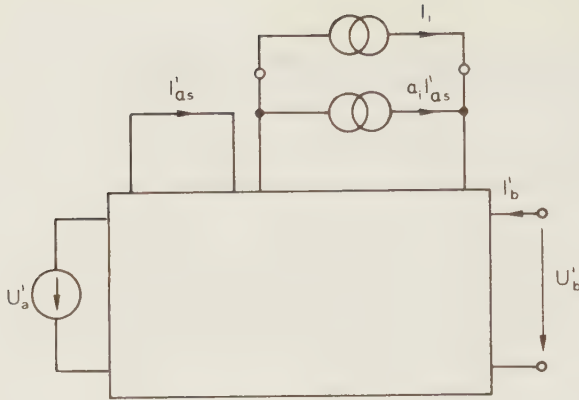


$$S_{g_i}^H = h_{ia} h_{bi}$$

$$h_{ia} = \left. \frac{U'_{gs}}{U'_a} \right|_{I'_i, I'_b = 0}$$

$$h_{bi} = \left. \frac{U'_b}{I'_i} \right|_{U'_a, I'_b = 0}$$

b) Spannungsgesteuerte Stromquelle



$$S_{ai}^{ii} = h_{ia} h_{bi}$$

$$h_{ia} = \frac{I'_{as}}{U'_a} \Big|_{I_i, I'_b = 0}$$

$$h_{bi} = \frac{U'_b}{I_i} \Big|_{U'_a, I'_b = 0}$$

c) Stromgesteuerte Stromquelle

Bild 1.5

Es wurde gezeigt, wie die Empfindlichkeiten bezüglich der Impedanz oder Admittanz von passiven Bauelementen aus dem Produkt von zwei Übertragungsfunktionen berechnet werden können. In der Regel werden jedoch die Empfindlichkeiten bezüglich der frequenzunabhängigen Parameter R , L , C (oder deren Kehrwerte) interessieren. Diese können dann aus den Empfindlichkeiten bezüglich der Impedanz bzw. der Admittanz durch Anwendung der Kettenregel der Ableitung berechnet werden. So ist, ausgehend von der Empfindlichkeit bezüglich der Impedanz

$$S_R^H = S_Z^H, \quad (1.48a)$$

$$S_L^H = p S_Z^H, \quad (1.48b)$$

$$S_C^H = -p Z^2 S_Z^H, \quad (1.48c)$$

oder ausgehend von der Empfindlichkeit bezüglich der Admittanz

$$S_R^H = -Y^2 S_Y^H, \quad (1.49a)$$

$$S_L^H = -p Y^2 S_Y^H, \quad (1.49b)$$

$$S_C^H = p S_Y^H. \quad (1.49c)$$

Daraus ist zu erkennen, daß es günstiger ist, die Empfindlichkeiten bezüglich der Widerstände und Induktivitäten aus den Empfindlichkeiten bezüglich deren Impedanz gemäß Gl. (1.48a) bzw. Gl. (1.48b) und die Empfindlichkeiten bezüglich der Kapazitäten aus den Empfindlichkeiten bezüglich deren Admittanz gemäß Gl. (1.49c) zu berechnen. Entsprechendes gilt für die Empfindlichkeiten bezüglich G , L^{-1} und C^{-1} .

In Bild 1.6 wird ein allgemeines Netzwerk gezeigt, in dem außer den passiven Bauelementen alle Arten von gesteuerten Quellen vorhanden sein können. Jede Bauelementeart steht hier stellvertretend für alle der gleichen Art, die im Netzwerk enthalten sind.

Auf Grund des Überlagerungssatzes kann folgendes Gleichungssystem für die Verknüpfung zwischen den abhängigen und den unabhängigen Größen in Bild 1.6 aufgestellt werden:

$$\begin{bmatrix} U_b \\ I_R \\ I_L \\ U_C \\ U_{\mu s} \\ I_{rs} \\ U_{gs} \\ I_{\alpha s} \end{bmatrix} = \begin{bmatrix} h_{ba} & h_{bR} & h_{bL} & h_{bC} & h_{b\mu} & h_{br} & h_{bg} & h_{b\alpha} \\ h_{Ra} & h_{RR} & h_{RL} & \dots & \dots & \dots & \dots & h_{R\alpha} \\ h_{La} & h_{LR} & \dots & \dots & \dots & \dots & \dots & \dots \\ h_{Ca} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ h_{\mu a} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ h_{ra} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ h_{ga} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ h_{\alpha a} & h_{\alpha R} & \dots & \dots & \dots & \dots & \dots & h_{\alpha \alpha} \end{bmatrix} \begin{bmatrix} U_a \\ U_R \\ U_L \\ I_C \\ U_{\mu} \\ U_r \\ I_g \\ I_{\alpha} \end{bmatrix} \quad (1.50)$$

Die Empfindlichkeiten der Leerlauf-Übertragungsfunktion $H = U_b/U_a$ bezüglich der Parameter R , μ , r , g und α werden direkt aus dem Produkt der entsprechenden Übertragungsfunktionen der ersten Zeile und der ersten Spalte der Übertragungsmatrix aus Gl. (1.50) gebildet. Für die Empfindlichkeiten bezüglich der Induktivitäten und Kapazitäten des Netzwerkes müssen die entsprechenden Produkte zusätzlich mit p (gemäß Gln. (1.48b) und (1.49c)) multipliziert werden.

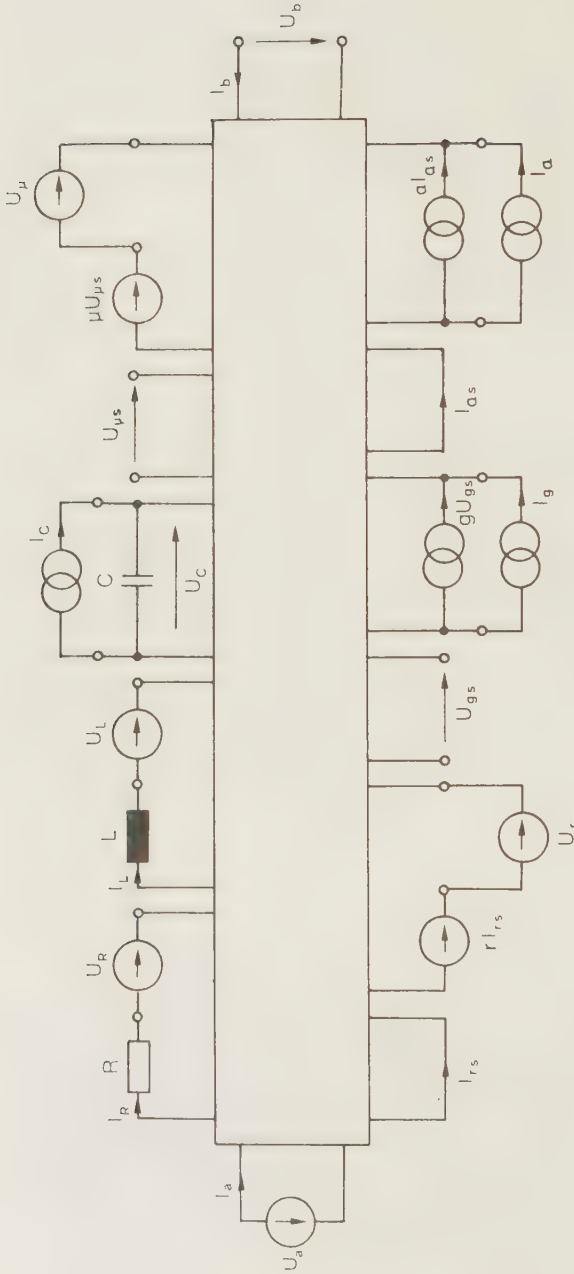


Bild 1.6: Allgemeines Netzwerk für die Empfindlichkeitsanalyse

2.2 Empfindlichkeiten zweiter Ordnung

Wie im folgenden gezeigt wird, können die Empfindlichkeiten zweiter Ordnung bezüglich der passiven oder aktiven Parameter ebenfalls als Produkte und Summen von Übertragungsfunktionen aus der Hybridmatrix in Gl. (1.50) ausgedrückt werden. Für den Fall der Empfindlichkeit zweiter Ordnung bezüglich einer Impedanz geht man von Gl. (1.39) aus. Durch Ableitung nach Z_j ist

$$\frac{\partial^2 H}{\partial Z_i \partial Z_j} = h_{ia} \frac{\partial h_{bi}}{\partial Z_j} + \frac{\partial h_{ia}}{\partial Z_j} h_{bi} \quad (1.51)$$

Hierin sind $\frac{\partial h_{ia}}{\partial Z_j}$ und $\frac{\partial h_{bi}}{\partial Z_j}$ die Empfindlichkeiten der Übertragungsfunktion h_{ia} und h_{bi} bezüglich der Impedanz Z_j . Diese lassen sich nun wie im Abschnitt 2.1.1 gezeigt, durch das Produkt von jeweils zwei Übertragungsfunktionen ausdrücken. Für die Empfindlichkeit von h_{ia} ist a als Eingangs- und i als Ausgangstor zu betrachten. Es wird daher die Übertragungsfunktion h_{ja} vom Eingang zum j -ten Element und die Übertragungsfunktion h_{ij} von diesem Element zum Ausgang benötigt. Daher ist

$$\frac{\partial h_{ia}}{\partial Z_j} = h_{ja} h_{ij} \quad (1.52a)$$

In ähnlicher Weise erhält man

$$\frac{\partial h_{bi}}{\partial Z_j} = h_{ji} h_{bj} \quad (1.52b)$$

Durch Einsetzen der Gln. (1.52a,b) in Gl. (1.51) ist schließlich

$$S_{Z_i Z_j}^H = \frac{\partial^2 H}{\partial Z_i \partial Z_j} = h_{ia} h_{ji} h_{bj} + h_{ja} h_{ij} h_{bi} \quad (1.53)$$

Die hierfür zusätzlich benötigten Übertragungsfunktionen zwischen den Elementen, nämlich h_{ij} und h_{ji} sind ebenfalls in der Hybridmatrix aus Gl. (1.28) enthalten.

In ähnlicher Weise kann gezeigt werden, daß auch die Empfindlichkeiten zweiter Ordnung bezüglich einer Admittanz oder des Streufaktors einer gesteuerten Quelle wie in Gl. (1.53) aus der Summe von Produkten spezieller Übertragungsfunktionen der Matrix aus

Gl. (1.50) berechnet werden können.

Die Empfindlichkeiten höherer Ordnung können durch Differenzieren der Gl. (1.53) gewonnen werden. Es treten dann auch weitere Übertragungsfunktionen auf, die ebenfalls in der Matrix aus Gl. (1.50) enthalten sind.

2.3 Die praktische Berechnung der Empfindlichkeiten

Die Matrix der Übertragungsfunktionen aus Gl. (1.50) läßt sich mit Hilfe des allgemeinen Netzwerkanalyseprogramms ERLNAP [15] in einem einzigen Analyseschritt berechnen. Dieses Programm arbeitet mit Zustandsvariablen und gibt die Übertragungsfunktionen in gebrochen rationaler Form als Funktionen des komplexen Frequenzparameters p an. Die nach diesem Verfahren mit Hilfe des o.g. Programmes berechneten Empfindlichkeiten erfüllen die im Abschnitt 1 dieses Kapitels aufgeführten Bedingungen (a) bis (c), d.h.:

- (a) Es werden die Empfindlichkeiten einer Netzwerkfunktion berechnet (hier wurde als Beispiel die Leerlauf-Übertragungsfunktion genommen);
- (b) Es können die Empfindlichkeiten erster und zweiter Ordnung gleichzeitig gegenüber sämtlichen passiven und aktiven Parametern mit einem einzigen Analyseschritt berechnet werden;
- (c) Da die Übertragungsfunktionen als rationale Funktionen in p berechnet werden, erhält man auch die Empfindlichkeiten als solche Funktionen.

3. Berechnung der Empfindlichkeiten durch differenzieren einer in symbolischer Form gegebenen Netzwerkfunktion

Für die Empfindlichkeitsberechnung in Optimierungsaufgaben ist das in Abschnitt 2 beschriebene Verfahren weniger geeignet, da in jedem Optimierungsschritt das Netzwerk für den geänderten Satz der Parameter neu analysiert werden müßte. Bei solchen Aufgabenstellungen ist es günstiger, die Empfindlichkeiten durch Ableiten der Netzwerkfunktion nach den Parametern zu berechnen. Die Netzwerkfunktion muß dazu in symbolischer Form vorliegen, d.h. die Koeffizienten der rationalen Funktion in p müssen als algebraische Ausdrücke in den Netzwerkparametern gegeben sein. Symbolische Netzwerkfunktionen können mit topologischen Netzwerkanalyseprogrammen generiert werden. Die Empfindlichkeiten erster und zweiter Ordnung werden daraus ebenfalls in symbolischer Form berechnet. Die Netzwerkfunktion und die Empfindlichkeiten in symbolischer Form braucht man daher nur einmal vor Beginn des Optimierungsprozesses zu erzeugen. Während der Optimierung kann dann für jeden beliebigen Satz von Parameterwerten und jeden beliebigen Frequenzpunkt sowohl die Netzwerkfunktion als auch deren Empfindlichkeiten numerisch berechnet werden.

3.1 Darstellung einer Netzwerkfunktion in symbolischer Form

Bekanntlich ist eine Netzwerkfunktion eines beliebigen linearen, zeitinvarianten und aus konzentrierten Bauelementen bestehenden Netzwerkes eine rationale Funktion in p von der Art

$$F(\underline{x}, p) = \frac{Z(\underline{x}, p)}{N(\underline{x}, p)} = \frac{\sum_{k=1}^r a_k(\underline{x}, p)}{\sum_{l=1}^s b_l(\underline{x}, p)} \quad . \quad (1.54)$$

Hierin ist $\underline{x} = (x_1, x_2, \dots, x_n)^t$ (1.55)

der Vektor der n Netzwerkparameter, r die Gesamtzahl der Zähler- und s die Gesamtzahl der Nennerterme. Die Terme $a_k(\underline{x}, p)$ und

$b_1(x, p)$ sind Vorzeichenbehaftete Produkte der Parameter $x_1, x_2, \dots, x_n$ und einer p -Potenz, wobei die Faktoren $x_1, x_2, \dots, x_n$ nur jeweils höchstens einmal vorkommen.*)

Sowohl für die automatische Berechnung der Netzwerkfunktion für einen beliebigen Parametervektor und einen beliebigen Frequenzpunkt als auch für die Generierung der Empfindlichkeiten in symbolischer Form auf dem Digitalrechner ist es zweckmäßig, Zähler und Nenner aus Gl. (1.54), ähnlich wie in [16] und [17] in jeweils einer Matrix mit folgenden Eigenschaften darzustellen:

- (a) Jeder Term entspricht einer Zeile der Matrix.
- (b) In der ersten Spalte steht der numerische Faktor des Terms.
- (c) Die nächsten n Spalten entsprechen jeweils einem Netzwerkparameter. Enthält der Term den Parameter x_i , so ist die Spalte $i+1$ mit einer 1 besetzt, sonst steht an dieser Stelle eine 0.
- (d) Die letzte Spalte enthält die Potenz von p des Terms.

Am besten kann diese Darstellungsweise mit Hilfe des folgenden Beispiels verdeutlicht werden.

Beispiel: Der aktive Tiefpaß zweiten Grades gemäß Bild 1.7 ([5], S. 365) hat die Übertragungsfunktion

$$H(p) = \frac{U_2}{U_1} = \frac{g_1 g_2 k_1 k_2}{C_1 C_2 p^2 + g_1 C_2 p + g_2 C_1 p - g_2 C_1 k_1 k_2 p + g_1 g_2} \quad (1.56)$$

*) Enthält das Netzwerk einen Übertrager mit dem Übertragungsfaktor $\ddot{u}$ oder einen Gyrator mit dem Gyrationswiderstand r , können die Koeffizienten a_k bzw. b_1 auch die Faktoren $\ddot{u}^2$ oder r^2 enthalten. Diese Schwierigkeit kann jedoch dadurch umgangen werden, daß $\ddot{u}^2$ durch $\ddot{u}_1 \cdot \ddot{u}_2$ bzw. r^2 durch $r_1 \cdot r_2$ ersetzt wird.

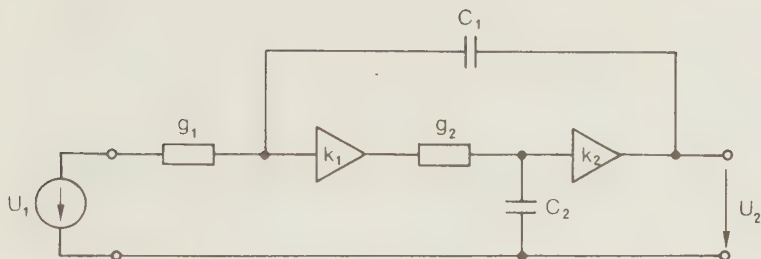


Bild 1.7: Aktiver Tiefpaß zweiten Grades

Werden die Parameter im Vektor

$$\underline{x} = (x_1, x_2, x_3, x_4, x_5, x_6)^t = (g_1, g_2, C_1, C_2, k_1, k_2)^t \quad (1.57)$$

zusammengefaßt, ist

$$H(\underline{x}, p) = \frac{Z(\underline{x}, p)}{N(\underline{x}, p)} = \frac{x_1 x_2 x_5 x_6}{x_3 x_4 p^2 + x_1 x_4 p + x_2 x_3 p - x_2 x_3 x_5 x_6 p + x_1 x_2} \quad (1.58)$$

die der Gl. (1.54) entsprechende Schreibweise der Übertragungsfunktion in symbolischer Form. Zähler und Nenner lassen sich daher nach den o.g. Kodierungsregeln folgendermaßen darstellen:

	Faktor	x_1	x_2	x_3	x_4	x_5	x_6	p
<u>Zähler:</u> Term a_1	1	1	1	0	0	1	1	0

<u>Nenner:</u> Term b_1		1	0	0	1	1	0	0	2
b_2		1	1	0	0	1	0	0	1
b_3		1	0	1	1	0	0	0	1
b_4		-1	0	1	1	0	1	1	1
b_5		1	1	1	0	0	0	0	0

(1.59)

In diesem Beispiel wurde die Übertragungsfunktion in "voll-symbolischer" Form dargestellt, d.h. die Werte aller Netzwerkelemente wurden durch Symbole ersetzt. Aus diesem Grund sind die numerischen Faktoren der Terme in der ersten Spalte der Matrizen immer gleich ± 1 (bei passiven Netzwerken können die Terme nur ein positives Vorzeichen haben). Bei vielen Anwendungen ist jedoch nur eine "teilsymbolische" Form der Netzwerkfunktion notwendig, d.h. es werden nur einige Bauelemente durch ihre Formelsymbole dargestellt benötigt. Da für die restlichen Elemente ihre numerischen Werte eingesetzt werden, haben die Faktoren in der ersten Spalte i.allg. einen von ± 1 unterschiedlichen reellen Zahlenwert.

3.2 Empfindlichkeiten erster Ordnung in symbolischer Form

Die vollrelative Empfindlichkeit der Netzwerkfunktion Gl. (1.54) kann gemäß Gl. (1.9) aus der vollrelativen Empfindlichkeit von Zähler und Nenner berechnet werden

$$S_{x_i}^{F(\underline{x},p)} = S_{x_i}^{Z(\underline{x},p)} - S_{x_i}^{N(\underline{x},p)} . \quad (1.60)$$

Da Zähler und Nenner jeden einzelnen Parameter nur in der ersten Potenz enthält, können deren Ableitungen nach x_i einfach aus der mit x_i dividierten Summe der Terme gebildet werden, die x_i als Faktor enthalten:

$$S_{x_i}^{Z(\underline{x},p)} = \frac{\partial Z(\underline{x},p)}{\partial x_i} = \frac{1}{x_i} \sum_{k \in K_i} a_k(\underline{x},p) \quad (1.61)$$

$$S_{x_i}^{N(\underline{x},p)} = \frac{\partial N(\underline{x},p)}{\partial x_i} = \frac{1}{x_i} \sum_{l \in L_i} b_l(\underline{x},p) . \quad (1.62)$$

Die vollrelativen Zähler- und Nennerempfindlichkeiten sind dann:

$$S_{x_i}^{Z(\underline{x},p)} = \frac{x_i}{Z(\underline{x},p)} \frac{\partial Z(\underline{x},p)}{\partial x_i} = \frac{1}{Z(\underline{x},p)} \sum_{k \in K_i} a_k(\underline{x},p) \quad (1.63)$$

$$S_{x_i}^{N(\underline{x}, p)} = \frac{x_i}{N(\underline{x}, p)} \frac{\partial N(\underline{x}, p)}{\partial x_i} = \frac{1}{N(\underline{x}, p)} \sum_{l \in L_i} b_l(\underline{x}, p) \quad (1.64)$$

Hierin ist

$$K_i = \{k \mid a_k(\underline{x}, p) = f(x_i)\} \quad (1.65)$$

die Menge aller Indizes der Zählerterme a_k und

$$L_i = \{l \mid b_l(\underline{x}, p) = f(x_i)\} \quad (1.66)$$

die Menge aller Indizes der Nennerterme b_l , die den Parameter x_i als Faktor enthalten.

Nach Einsetzen der Gln. (1.63) und (1.64) in Gl. (1.60) ist

$$S_{x_i}^{F(\underline{x}, p)} = \frac{1}{Z(\underline{x}, p)} \sum_{k \in K_i} a_k(\underline{x}, p) - \frac{1}{N(\underline{x}, p)} \sum_{l \in L_i} b_l(\underline{x}, p) \quad (1.67)$$

Die Berechnung der Empfindlichkeiten erster Ordnung in symbolischer Form reduziert sich daher auf die Erzeugung der Indexmengen K_i und L_i für $i=1,2,\dots,n$. Diese können nun sehr leicht aus der Zähler- bzw. Nennermatrix der Netzwerkfunktion in symbolischer Form generiert werden. Hierzu werden in die Indexmenge K_i bzw. L_i diejenigen Zeilen eingetragen, in denen in der Spalte x_i eine 1 steht. Diese Indexmengen brauchen nur einmal für jedes Netzwerk erzeugt zu werden. Für jeden beliebigen Parametervektor $\underline{x}$ und an jedem beliebigen Frequenzpunkt $p = j\omega$ können dann die Empfindlichkeiten mit Hilfe der Gl. (1.67) mit geringem Aufwand berechnet werden.

Beispiel: Es sind die vollrelativen Empfindlichkeiten erster Ordnung der Übertragungsfunktion bezüglich aller 6 Parameter des Beispiels aus Abschnitt 3.1 zu berechnen.

Zunächst werden hierzu gemäß Gln.(1.65) und (1.66) aus dem Schema (1.59) folgende Indexmengen gebildet:

$$K_1 = (1), K_2 = (1), K_3 = (), K_4 = (), K_5 = (1), K_6 = (1)$$

$$\begin{aligned} L_1 &= (2, 5), & L_2 &= (3, 4, 5) & L_3 &= (1, 3, 4), \\ L_4 &= (1, 2), & L_5 &= (4), & L_6 &= (4). \end{aligned}$$

Die Indextmengen K_3 und K_4 sind leer, da der Zähler der Übertragungsfunktion, Gl. (1.58), nicht von x_3 und x_4 abhängt. Mit diesen Indextmengen können nun gemäß Gl. (1.67) die Empfindlichkeiten der Übertragungsfunktion bezüglich der Bauelemente berechnet werden. Man erhält:

$$\begin{aligned} S_{x_1}^H &= \frac{1}{Z} a_1 - \frac{1}{N} (b_2 + b_5) \\ S_{x_2}^H &= \frac{1}{Z} a_1 - \frac{1}{N} (b_3 + b_4 + b_5) \\ S_{x_3}^H &= -\frac{1}{N} (b_1 + b_3 + b_4) \\ S_{x_4}^H &= -\frac{1}{N} (b_1 + b_2) \\ S_{x_5}^H &= \frac{1}{Z} a_1 - \frac{1}{N} b_4 \\ S_{x_6}^H &= \frac{1}{Z} a_1 - \frac{1}{N} b_4. \end{aligned}$$

Bisher wurde nur gezeigt, wie die vollrelative Empfindlichkeit erster Ordnung in symbolischer Form berechnet werden kann. Die absolute und die halbrelative Empfindlichkeit lassen sich mit Hilfe der Gln. (1.2) und (1.3) sofort aus Gl. (1.67) angeben.

$$S_{x_i}^F(\underline{x}, p) = \frac{1}{x_i N(\underline{x}, p)} \sum_{k \in K_i} a_k(\underline{x}, p) - \frac{Z(\underline{x}, p)}{x_i N^2(\underline{x}, p)} \sum_{l \in L_i} b_l(\underline{x}, p) \quad (1.68)$$

$$S_{x_i}^F(\underline{x}, p) = \frac{1}{N(\underline{x}, p)} \sum_{k \in K_i} a_k(\underline{x}, p) - \frac{Z(\underline{x}, p)}{N^2(\underline{x}, p)} \sum_{l \in L_i} b_l(\underline{x}, p) \quad (1.69)$$

3.3 Empfindlichkeiten zweiter Ordnung in symbolischer Form

Es soll hier nur die absolute Empfindlichkeit zweiter Ordnung betrachtet werden, da sich die halb- und vollrelative aus der ersten nach Gln. (1.24a,b) sofort angeben lassen. Die absolute Empfindlichkeit zweiter Ordnung ist

$$\begin{aligned}
 S_{x_i x_j}^F(\underline{x}, p) &= \frac{\partial^2 F(\underline{x}, p)}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_j} \left(\frac{\partial}{\partial x_i} \left(\frac{Z(\underline{x}, p)}{N(\underline{x}, p)} \right) \right) & (1.70) \\
 &= \frac{\partial}{\partial x_j} \left(\frac{1}{N} \cdot \frac{\partial Z}{\partial x_i} - \frac{Z}{N^2} \frac{\partial N}{\partial x_i} \right) \\
 &= -\frac{1}{N^2} \frac{\partial N}{\partial x_j} \frac{\partial Z}{\partial x_i} + \frac{1}{N} \frac{\partial^2 Z}{\partial x_i \partial x_j} - \left(\frac{1}{N^2} \frac{\partial Z}{\partial x_j} - 2 \frac{Z}{N^3} \frac{\partial N}{\partial x_j} \right) \frac{\partial N}{\partial x_i} - \frac{Z}{N^2} \frac{\partial^2 N}{\partial x_i \partial x_j} \\
 S_{x_i x_j}^F &= \frac{Z}{N} \left\{ -\frac{1}{ZN} (S_{x_i}^Z S_{x_j}^N + S_{x_j}^Z S_{x_i}^N) + \frac{2}{N^2} S_{x_i}^N S_{x_j}^N + \frac{1}{Z} S_{x_i x_j}^Z - \frac{1}{N} S_{x_i x_j}^N \right\}. & (1.71)
 \end{aligned}$$

Die Zähler- und Nennerempfindlichkeiten werden gemäß Gln. (1.61) und (1.62) berechnet. In Gl. (1.71) müssen daher nur noch die Empfindlichkeiten zweiter Ordnung von Zähler und Nenner berechnet werden. Dies geschieht analog zum Fall der Empfindlichkeiten erster Ordnung. Für die Ableitungen zweiter Ordnung brauchen nur solche Terme berücksichtigt werden, in denen das Produkt $x_i x_j$ enthalten ist. Es ist also:

$$S_{x_i x_j}^Z(\underline{x}, p) = \frac{\partial^2 Z(\underline{x}, p)}{\partial x_i \partial x_j} = \frac{1}{x_i x_j} \sum_{k \in K_{ij}} a_k(\underline{x}, p) \quad (1.72)$$

und

$$S_{x_i x_j}^N(\underline{x}, p) = \frac{\partial^2 N(\underline{x}, p)}{\partial x_i \partial x_j} = \frac{1}{x_i x_j} \sum_{l \in L_{ij}} b_l(\underline{x}, p) \quad (1.73)$$

Hierin ist

$$K_{ij} = \{k \mid a_k(\underline{x}, p) = f(x_i x_j)\} \quad (1.74)$$

die Menge aller Indizes der Zählerterme a_k und

$$L_{ij} = \{1 | b_1(\underline{x}, p) = f(x_i x_j)\} \quad (1.75)$$

die Menge aller Indizes der Nennerterme b_1 , die das Produkt $x_i x_j$ als Faktor enthalten. Im Prinzip müßten für die Empfindlichkeitsberechnung zweiter Ordnung nach Gln. (1.72) und (1.73) insgesamt $2n^2$ Indexmengen gebildet werden. Da jedoch Zähler und Nenner jeweils die Parameter nur in ihrer ersten Potenz enthalten, gilt

$$S_{x_i x_i}^Z = S_{x_i x_i}^N = 0 \quad \text{für } i = 1, \dots, n. \quad (1.76)$$

Da außerdem die Reihenfolge der Ableitung unbedeutend ist, also

$$S_{x_i x_j}^Z = S_{x_j x_i}^Z \quad \text{und} \quad S_{x_i x_j}^N = S_{x_j x_i}^N \quad (1.77)$$

gilt, werden insgesamt nur $n(n-1)$ Indexmengen benötigt.

Beispiel: Es sind die absoluten Empfindlichkeiten zweiter Ordnung der Übertragungsfunktion des Beispiels aus 3.1 zu berechnen.

Nach Gl. (1.71) sind zunächst die Zähler und Nennerempfindlichkeiten zu berechnen. Diese erhält man nach Gl. (1.61) und (1.62), wie im vorangegangenen Abschnitt gezeigt wurde. Für die Berechnung der Empfindlichkeiten zweiter Ordnung von Zähler und Nenner sind erst die Indexmengen K_{ij} und L_{ij} für $i = 1, 2, \dots, n$ und $j > i$ zu bilden. Mit der Vorschrift aus den Gln. (1.74) und (1.75) werden diese direkt aus dem Schema (1.59) gewonnen, indem diejenigen Zeilen in die Indexmenge eingetragen werden, in denen gleichzeitig in den Spalten x_i und x_j eine 1 steht. Diese Indexmengen werden hier zwecks besserer Übersicht in die beiden Matrizen $\underline{K} = [K_{ij}]$ und $\underline{L} = [L_{ij}]$ eingetragen.

$$\underline{K} = \begin{bmatrix} (1) & () & () & (1) & (1) \\ & () & () & (1) & (1) \\ & & () & () & () \\ & & & () & () \\ & & & & (1) \end{bmatrix} \quad (1.78)$$

$$\underline{L} = \begin{bmatrix} & (5) & () & (2) & () & () \\ & & (3,4) & () & (4) & (4) \\ & & & (1) & (4) & (4) \\ & & & & () & () \\ & & & & & (4) \end{bmatrix} \quad (1.79)$$

Die Indexmengen auf den Hauptdiagonalen sind wegen Gl. (1.76) alle leer. Beide Matrizen sind wegen Gl. (1.77) bezüglich ihrer Hauptdiagonalen symmetrisch.

Aus diesen Indexmengen können nun die Zähler- und Nennerempfindlichkeiten zweiter Ordnung gemäß Gln. (1.72) und (1.73) berechnet und in Gl. (1.71) eingesetzt werden. Als Beispiel soll jedoch nur die Empfindlichkeit zweiter Ordnung bezüglich x_3 und x_4 vollständig berechnet werden.

Mit den Indexmengen K_3 , K_4 , L_3 und L_4 erhält man auf Grund der Gln. (1.61) und (1.62)

$$S_{x_3}^Z = 0, \quad S_{x_4}^Z = 0$$

$$S_{x_3}^N = \frac{1}{x_3}(b_1+b_3+b_4), \quad S_{x_4}^N = \frac{1}{x_4}(b_1+b_2).$$

Mit den Indexmengen K_{34} und L_{34} aus den Gln. (1.78) und (1.79) erhält man auf Grund der Gln. (1.72) und (1.73)

$$S_{x_3 x_4}^Z = 0, \quad S_{x_3 x_4}^N = \frac{1}{x_3 x_4} b_1.$$

Werden diese Werte in Gl. (1.71) eingeführt, erhält man

$$S_{x_3 x_4}^H = \frac{Z}{N x_3 x_4} \left\{ \frac{2}{N^2} (b_1+b_3+b_4)(b_1+b_2) - \frac{b_1}{N} \right\}.$$

Wie leicht nachgeprüft werden kann, ist dies das gleiche Ergebnis, das durch zweifache Ableitung nach x_3 und x_4 der Übertragungsfunktion aus Gl. (1.58) entsteht.

II. DIE EMPFINDLICHKEITSOPTIMIERUNG LINEARER NETZWERKE

1. Einführung

Im Kapitel I wurden zwei verschiedene Methoden für die Berechnung der Empfindlichkeiten einer Netzwerkfunktion bezüglich der Netzwerkparameter beschrieben. Auf Grund der Einzelempfindlichkeiten kann ein Maß für die Gesamtempfindlichkeit des Netzwerkes berechnet werden. Mit Hilfe eines solchen Maßes können verschiedene Schaltungen, die eine gleiche Netzwerkfunktion realisieren, in Bezug auf ihre Gesamtempfindlichkeit miteinander verglichen werden. Um bei einem Vergleich dieser Art zu einer objektiven Aussage zu gelangen, ist es notwendig, jede der Vergleichsschaltungen erst so zu dimensionieren, daß sie bei fest vorgegebener Struktur und zu realisierender Netzwerkfunktion ihre minimal erreichbare Gesamtempfindlichkeit besitzt. Dies ist natürlich nur dann möglich, wenn bei der Auswahl der Netzwerkparameter Freiheitsgrade vorhanden sind, d.h. wenn einige von ihnen frei gewählt werden können.

Werden die vorhandenen Freiheitsgrade bei der Dimensionierung der zu vergleichenden Netzwerke nicht genutzt um eine minimale Gesamtempfindlichkeit zu erreichen, kann man zu falschen Ergebnissen kommen. So kann z.B. durch die zufällige Auswahl der freiwählbaren Parameter ein Netzwerk A eine geringere Gesamtempfindlichkeit als ein äquivalentes Netzwerk B besitzen, obwohl nach einer Optimierung das Verhältnis umgekehrt wäre, also B die geringere Gesamtempfindlichkeit hätte.

Die Bestimmung der Netzwerkparameter für die minimale Gesamtempfindlichkeit geschieht mit Hilfe nichtlinearer Programmierungsmethoden (Optimierung mit Nebenbedingungen), wie sie im Abschnitt 3 beschrieben werden. Zunächst ist jedoch ein geeignetes Maß für die Gesamtempfindlichkeit eines Netzwerkes aufzustellen.

2. Das Empfindlichkeitsmaß

2.1 Vergleich der Empfindlichkeitsmaße

In der Literatur sind verschiedene Maße für die Beurteilung der Gesamtempfindlichkeit eines Netzwerkes vorgeschlagen worden [7], [18], [19], [20], [21], [22], [23]. Diese können in zwei Kategorien aufgeteilt werden. Zu der ersten Kategorie gehören solche, die als "reine" Empfindlichkeitsmaße bezeichnet werden können, da sie allein aus den Einzelempfindlichkeiten bezüglich aller Parameter berechnet werden. Hierzu zählt das Empfindlichkeitsmaß von Goldstein und Kuo [7] und das von Schoeffler [20]. Zu der zweiten Kategorie gehören solche, die aus der Abweichung der Netzwerkfunktion von ihrem Sollwert, bei bekannten Toleranzen der Bauelemente, ermittelt werden. Die Abweichung wird dabei entweder genau wie von Hakimi und Cruz [18] und Kelly [19] oder nur näherungsweise (mit Hilfe der Empfindlichkeiten) wie von Rosenblum und Ghausi [21], Shenoi [22], und Haykin und Butler [23] berechnet. In jedem Fall werden aber die Toleranzen der Bauelemente oder deren Mittelwert und Streuung benötigt. Meistens werden diese Maße ebenfalls einfach Empfindlichkeitsmaße genannt. Da sie jedoch auf Grund von Abweichungen der Netzwerkfunktion von ihrem Sollwert berechnet werden, d.h. auf einem Fehler basieren, sollen sie hier genauer mit "Fehler"- oder "Toleranzmaße" bezeichnet werden.

Es ergibt sich nun die Frage, welches Maß für die Empfindlichkeitsoptimierung herangezogen werden soll. Um diese Frage beantworten zu können, müssen zwei verschiedene Anwendungsfälle betrachtet werden.

- (a) Empfindlichkeitsoptimierung einer Schaltung, in der die Bauelementetoleranzen noch nicht festliegen. Dieser Fall tritt dann auf, wenn ein grundsätzlicher Vergleich verschiedener Schaltungen mit der gleichen Netzwerkfunktion durchgeführt werden soll, um die in Bezug auf Gesamtempfindlichkeit günstigste unter ihnen auszusuchen. Die statistischen Kenngrößen der Bauelemente (Toleranz, Mittelwert und Varianz) sind zu diesem Zeitpunkt in der Regel noch gar nicht bekannt, da sie

meistens erst später, z.B. durch eine Kostenoptimierung, festgelegt werden. Aus diesem Grund kann hier als Empfindlichkeitsmaß nur eines der ersten Kategorie, d.h. ein "reines" Empfindlichkeitsmaß, verwendet werden.

- (b) Eine Schaltung, in der die statistischen Kenngrößen der Bauelemente bekannt sind, soll bezüglich ihrer Gesamtempfindlichkeit optimiert werden. Dieser Fall ist dann gegeben, wenn Netzwerke verglichen werden sollen, deren Bauelementetoleranzen aus irgendwelchen Gründen von Anfang an festliegen. Dieser Fall würde auch dann auftreten, wenn es für eine vorgegebene Netzwerkfunktion nur eine Realisierungsmöglichkeit gäbe. Es wäre dann nämlich sinnvoll, die Bauelementetoleranzen bereits vor der Empfindlichkeitsoptimierung festzulegen.

In diesem Anwendungsfall wird natürlich ein Maß der zweiten Kategorie, d.h. ein Fehlermaß, zu bevorzugen sein, da bei Kenntnis der statistischen Größen ein solches Maß viel genauer die tatsächlichen Abweichungen der Netzwerkfunktion vom Sollwert widerspiegelt. Da es sich dabei eigentlich um ein Fehler- statt um ein reines Empfindlichkeitsmaß handelt, wäre die Optimierung hier eher als "Fehler"- oder "Toleranz"- statt Empfindlichkeitsoptimierung zu bezeichnen.

Im folgenden soll nur noch der Fall (a), nämlich die reine Empfindlichkeitsoptimierung behandelt werden. Fall (b) kann dann in völlig analoger Weise behandelt werden, wenn an Stelle des Empfindlichkeitsmaßes ein Fehlermaß eingesetzt wird.

Für die weiteren Betrachtungen kommt daher nur noch das Empfindlichkeitsmaß von Goldstein und Kuo [7] oder das von Schoeffler [20] in Frage.

Goldstein und Kuo hat als Maß für die Gesamtempfindlichkeit eines Netzwerkes an einem Frequenzpunkt $p = j\omega$ den Betrag des Empfindlichkeitsvektors aus Gl. (1.19) vorgeschlagen. Werden die halbrelativen Empfindlichkeiten verwendet, dann ist

$$\underline{S}^F(\underline{x}, j\omega) = \{S_{x_1}^F(\underline{x}, j\omega), S_{x_2}^F(\underline{x}, j\omega), \dots, S_{x_n}^F(\underline{x}, j\omega)\} \quad (2.1)$$

der Empfindlichkeitsvektor der (komplex- oder reellwertigen) Netzwerkfunktion $F(\underline{x}, j\omega)$ und

$$M(\omega, \underline{x}) = |S^F(\underline{x}, j\omega)| = \sqrt{\sum_{i=1}^n |S_{x_i}^F(\underline{x}, j\omega)|^2} \quad (2.2)$$

die Gesamtempfindlichkeit im Frequenzpunkt $p = j\omega$.

Von Schoeffler wurde als Maß für die Gesamtempfindlichkeit der quadrierte Wert von Gl. (2.2) vorgeschlagen [20]

$$M(\omega, \underline{x}) = \sum_{i=1}^n |S_{x_i}^F(\underline{x}, j\omega)|^2. \quad (2.3)$$

Für die Empfindlichkeitsoptimierung muß das Empfindlichkeitsmaß, und damit auch die Einzelempfindlichkeiten, als Funktion der Parameter x_i ($i=1, 2, \dots, n$) explizit gegeben sein. Aus diesem Grund sind in den Gln. (2.1) bis (2.3) die Abhängigkeiten sowohl von der Frequenz ω als auch vom Parametervektor $\underline{x} = (x_1, x_2, \dots, x_n)^T$ ausdrücklich angegeben worden. Um die Ausdrücke im folgenden nicht unnötig zu komplizieren und sie überschaubar zu halten, werden diese Abhängigkeitsangaben nicht immer mitgeführt.

Der Vollständigkeit halber soll anschließend gezeigt werden, daß die Empfindlichkeitsmaße gemäß Gln. (2.2) und (2.3) als Sonderfälle der Varianz bzw. der Dispersion einer reellen Netzwerkfunktion aufgefaßt werden können. Man betrachte dazu die relativen Änderungen der Parameter $\Delta x_i / x_i$ als voneinander unabhängige Zufallsvariablen mit dem Mittelwert 0 und der Varianz $\sigma_{x_i}^2$. Die Änderung der Netzwerkfunktion wird damit ebenfalls zur Zufallsvariable, deren Mittelwert und Varianz mit Hilfe der Gln. (A.5) und (A.12) aus dem Anhang berechnet werden können. Werden diese Gln. auf die linearisierte Änderung der Netzwerkfunktion

$$\Delta f = \sum_{i=1}^n S_{x_i}^f \frac{\Delta x_i}{x_i} \quad (2.4)$$

angewendet, erhält man

$$m_1 = E(\Delta F) = 0 \quad (2.5)$$

und

$$\sigma_{\Delta F}^2 = \sum_{i=1}^n (S_{x_i}^f)^2 \sigma_{x_i}^2. \quad (2.6)$$

Für den Sonderfall, daß alle Bauelemente die gleiche Varianz $\sigma_{x_i}^2 = 1$ ($i=1,2,\dots,n$) besitzen, stimmt Gl. (2.6) mit der Schoefflerschen Gesamtempfindlichkeit nach Gl. (2.3) einer reellen Netzwerkfunktion überein. Entsprechendes gilt für die Quadratwurzel aus Gl. (2.6), d.h. für die Dispersion $\sigma_{\Delta F}$, und das Empfindlichkeitsmaß nach Goldstein und Kuo gemäß Gl. (2.2).

Im weiteren Verlauf dieser Arbeit soll nur noch die Gesamtempfindlichkeit nach Schoeffler verwendet werden, da bei ihr gegenüber der Gesamtempfindlichkeit nach Goldstein und Kuo die Wurzelbildung wegfällt. Diese Tatsache ist bei der Verwendung des Empfindlichkeitsmaßes als Zielfunktion in der Empfindlichkeitsoptimierung von Bedeutung. Es müssen hierfür nämlich die Ableitungen eines Empfindlichkeitsmaßes für einen gesamten Frequenzbereich bezüglich der Netzwerkparameter gebildet werden, wobei man mit Gl. (2.3) zu einfacheren Ausdrücken als mit Gl. (2.2) kommt.

2.2 Ein integrales Empfindlichkeitsmaß

Gl. (2.3) ist ein Maß für die Gesamtempfindlichkeit an einem einzigen Frequenzpunkt. Wenn das Empfindlichkeitsmaß das Empfindlichkeitsverhalten eines Netzwerkes in einem Frequenzbereich $\omega \in (\omega_a, \omega_b)$ beschreiben soll, ist es naheliegend, Gl. (2.3) über diesen Frequenzbereich zu integrieren:

$$M(\underline{x}) = \int_{\omega_a}^{\omega_b} \left\{ \sum_{i=1}^n |S_{x_i}^F(\underline{x}, j\omega)|^2 \right\} d\omega. \quad (2.7)$$

Da im Rahmen dieser Arbeit die Empfindlichkeit eines Netzwerkes nicht mit einem einzelnen Frequenzpunkt, sondern über einem ge-

samen Frequenzintervall minimiert werden soll, wird von nun an nur noch das integrale Empfindlichkeitsmaß aus Gl. (2.7) verwendet werden.

2.3 Empfindlichkeitsmaß und Netzwerkfunktion

An dieser Stelle soll noch einiges über die Wahl der Netzwerkfunktion F in Gl. (2.7) gesagt werden. Die Wahl der Netzwerkfunktion, die das Netzwerk beschreibt, dessen Gesamtempfindlichkeit zu optimieren ist, hängt davon ab, welche Vorschrift das Netzwerk zu realisieren hat. Ist zum Beispiel ein vorgeschriebener Dämpfungsverlauf zu approximieren, so interessiert die Minimierung der Gesamtempfindlichkeit der Dämpfung. An Stelle von F ist dann in Gl. (2.7) z.B. die Dämpfung in Neper $a(\omega) = -\ln |H(j\omega)|$ einzusetzen. Ist dagegen ein vorgeschriebener Phasengang zu approximieren, so ist in Gl. (2.7) die Phase $\phi(\omega) = \arg H(j\omega)$ einzusetzen. Obwohl diese Forderung als selbstverständlich erscheint, ist sie bisher bei Gesamtempfindlichkeitsuntersuchungen und bei Empfindlichkeitsoptimierungen häufig nicht beachtet worden. So wird z.B. in [21] und [24] die Gesamtempfindlichkeit einer Übertragungsfunktion $H(j\omega)$ berechnet bzw. optimiert. Dies wird im allgemeinen zu falschen Ergebnissen führen, wenn das zu optimierende Netzwerk entweder nur einen vorgeschriebenen Dämpfungsverlauf oder nur einen vorgeschriebenen Phasengang zu approximieren hat, wie das meistens der Fall ist. Diese Tatsache soll im folgenden gezeigt werden.

Wird die Übertragungsfunktion $H(j\omega)$ in Abhängigkeit von Betrag und Phase ausgedrückt, dann ist

$$H(j\omega) = |H(j\omega)| e^{j\phi(\omega)} . \quad (2.8)$$

Mit der Dämpfung in Neper,

$$a(\omega) = -\ln |H(j\omega)| , \quad (2.9)$$

wird aus Gl. (2.8)

$$H(j\omega) = e^{-a(\omega) + j\phi(\omega)} . \quad (2.10)$$

Die halbrelative Empfindlichkeit der Übertragungsfunktion bezüglich eines Parameters x_i ist

$$S'_{x_i}^H(j\omega) = H(j\omega) \left(-S'_{x_i}^a(\omega) + j S'_{x_i}^\phi(\omega) \right) . \quad (2.11)$$

Hierin ist

$$S'_{x_i}^a(\omega) = \frac{\partial a(\omega)}{\partial x_i / x_i} \quad (2.12)$$

die halbrelative Empfindlichkeit der Dämpfung und

$$S'_{x_i}^\phi(\omega) = \frac{\partial \phi(\omega)}{\partial x_i / x_i} \quad (2.13)$$

die halbrelative Empfindlichkeit der Phase in Bezug auf den Parameter x_i .

Wird Gl. (2.11) in Gl. (2.7) eingeführt, ist das integrale Empfindlichkeitsmaß

$$M(\underline{x}) = \int_a^b |H(j\omega)|^2 \sum_{i=1}^n \left((S'_{x_i}^a(\omega))^2 + (S'_{x_i}^\phi(\omega))^2 \right) d\omega . \quad (2.14)$$

Das Empfindlichkeitsmaß ist also eine mit dem Faktor $|H(j\omega)|^2$ gewichtete Summe der Empfindlichkeiten der Dämpfung und der Phase. Dieser Umstand kann dazu führen, daß nach der Optimierung die Gesamtempfindlichkeit der Übertragungsfunktion zwar minimiert wurde, aber die Gesamtempfindlichkeit der Dämpfung (oder der Phase) größer ist als zuvor. Wenn die Schaltung aber einen vorgeschriebenen Verlauf der Dämpfung (oder der Phase) zu approximieren hat, wäre dieses Ergebnis unerwünscht. Gl. (2.14) ist daher nur in den selteneren Fällen anzuwenden, in denen gleichzeitig eine Dämpfungs- und eine Phasenvorschrift zu erfüllen ist.

3. Die Optimierung

3.1 Formulierung des Optimierungsproblems [24], [25]

Das Optimierungsproblem kann folgendermaßen formuliert werden:

An einem Netzwerk mit fest vorgegebener Struktur und bekannten Anfangswerten der Netzwerkelemente sind neue Elementewerte so zu bestimmen, daß das Empfindlichkeitsmaß zu einem Minimum wird. Es sind dabei bestimmte Nebenbedingungen für die Netzwerkfunktion und für die Parameter zu erfüllen.

Für die Netzwerkfunktion existieren dabei folgende zwei mögliche Forderungen [24] :

- (a) Wird das Netzwerk durch eine numerisch vorgegebene rationale Netzwerkfunktion beschrieben, so lassen sich die Nebenbedingungen in Form von p Gleichheitsbedingungen zwischen den realisierten und gewünschten Koeffizienten der Netzwerkfunktion ausdrücken

$$a_i(\underline{x}) = c_i \quad (i=1,2,\dots,p) . \quad (2.15)$$

Hierin ist $a_i(\underline{x})$ der i -te realisierte Koeffizient, dessen Abhängigkeit von den Netzwerkparametern bekannt sein muß, und c_i der i -te gewünschte Koeffizient im Zähler bzw. Nenner der Netzwerkfunktion.

- (b) Wird das Netzwerkverhalten durch ein Toleranzschema einer reellwertigen Netzwerkfunktion beschrieben, so lassen sich diese Anforderungen durch einen Satz von q frequenzabhängigen Ungleichheitsbedingungen formulieren

$$f_{iu} \leq f_i(\underline{x}) \leq f_{io} \quad (i=1,2,\dots,q) . \quad (2.16)$$

Hier ist $f_i(\underline{x})$ die realisierte reelle Netzwerkfunktion im i -ten Frequenzpunkt und f_{iu} bzw. f_{io} die durch das Toleranzschema gegebene untere bzw. obere Grenze in diesem Frequenzpunkt.

Für die Netzwerkparameter sind als Nebenbedingungen die durch ihre technische Realisierbarkeit gegebenen Grenzen zu beachten. Diese lassen sich durch die n Ungleichheitsbedingungen

$$x_{iu} \leq x_i \leq x_{io} \quad (i=1,2,\dots,n) \quad (2.17)$$

ausdrücken. In Gl. (2.17) bedeutet x_{iu} die untere und x_{io} die obere Realisierbarkeitsgrenze des Elements x_i .

Dieses Optimierungsproblem mit Nebenbedingungen, auch nichtlineares Programmierungsproblem (NLP) genannt, läßt sich für den allgemeinen Fall wie folgt formulieren:

Es ist der Vektor $\underline{x}^*$ der Netzwerkparameter so zu bestimmen, daß die Zielfunktion $Z(\underline{x})$ unter Berücksichtigung der m Ungleichheitsbedingungen

$$g_i(\underline{x}) \geq 0 \quad (i=1,2,\dots,m) \quad (2.18)$$

und der p Gleichheitsbedingungen

$$h_j(\underline{x}) = 0 \quad (j=1,2,\dots,p) \quad (2.19)$$

zu einem Minimum gemacht wird.

Bei der hier vorliegenden Aufgabe einer Empfindlichkeitsoptimierung dient das Empfindlichkeitsmaß nach Gl. (2.7) als Zielfunktion $Z(\underline{x})$.

3.2 Die modifizierte Zielfunktion

Für die Lösung des oben beschriebenen Optimierungsproblems mit Nebenbedingungen sind verschiedene Verfahren vorgeschlagen worden (siehe z.B. [26]). Eines davon ist die ursprünglich von Carroll [27] vorgeschlagene und später von Fiacco und McCormick [28] weiterentwickelte und mit "Sequential Unconstrained Minimization Technique" (SUMT) bezeichnete Methode. Hier soll eine Mo-

difikation dieser Methode eingesetzt werden.

Die Methode von Fiacco und McCormick basiert auf einer geänderten Zielfunktion, die außer der eigentlichen Zielfunktion $Z(\underline{x})$ zwei weitere Terme enthält:

$$P(\underline{x}, r) = Z(\underline{x}) + r_k \sum_{i=1}^m \frac{1}{g_i(\underline{x})} + \frac{1}{\sqrt{r_k}} \sum_{j=1}^p h_j^2(\underline{x}) . \quad (2.20)$$

Diese zusätzlichen Terme, auch Straffunktionen genannt, beeinflussen die modifizierte Zielfunktion in der Art, daß zunehmende Verletzungen der Nebenbedingungen steigende Werte der Zielfunktion $P(\underline{x}, r)$ ergeben.

Die Gewichtungsfaktoren r_k in Gl. (2.20) bilden eine monoton abnehmende Folge von positiven Werten. Das Optimierungsproblem mit Nebenbedingungen wird damit in eine Folge von unbeschränkten Optimierungen transformiert, deren Lösung im Grenzfall $r_k \rightarrow 0$ das Minimum des beschränkten Problems annähert. Der Vorteil dieser Methode liegt darin, daß das nicht beschränkte Optimierungsproblem mit einem der zahlreich hierfür existierenden Verfahren gelöst werden kann.

Die direkte Anwendung eines allgemeinen (unbeschränkten) Optimierungsverfahrens auf die modifizierte Zielfunktion aus Gl. (2.20), wie es z.B. in [24] und [25] getan wird, hat jedoch folgende Nachteile:

- Der Anfangswert $\underline{x}_0$ des Parametervektors muß so gewählt werden, daß alle Ungleichheitsbedingungen Gl. (2.18) erfüllt sind.
- Im Verlauf der Optimierung muß nach jedem Schritt nachgeprüft werden, ob alle Ungleichheitsbedingungen erfüllt sind. Ist dies nicht der Fall, muß die Schrittweite kürzer gewählt werden.

Werden einige der Ungleichheitsbedingungen am Anfang oder während der Optimierung nicht erfüllt, dann treten in der Summe im zweiten Term auf der rechten Seite von Gl. (2.20) negative Werte

auf, die eine ungewollte Reduktion der modifizierten Zielfunktion vortäuschen.

Um den hierdurch erforderlichen Eingriff in das unbeschränkte Optimierungsverfahren zu vermeiden, wird hier eine Variablentransformation eingeführt, mit der die Wahl eines geeigneten Startvektors $\underline{x}_0$ und die laufende Überprüfung der Ungleichheitsbedingungen während der Optimierung wegfällt. Diese Variablentransformation soll für den hier zu behandelnden Fall (a) aus 3.1 anwendbar sein. Werden nämlich die Forderungen an das Netzwerk durch einen Koeffizientenvergleich gemäß Gl. (2.15) als Gleichheitsbedingungen in die modifizierte Zielfunktion Gl. (2.20) eingebracht, dann bestehen die Ungleichheitsbedingungen lediglich aus Forderungen an die Netzwerkparameter nach Gl. (2.17).

Die Variablentransformation ist für den Fall, daß nur die Positivität der Parameter ($x_i \geq 0$ für $i = 1, 2, \dots, n$) verlangt wird [26]

$$x_i = y_i^2 \quad (2.21)$$

und für den Fall, daß die Parameter innerhalb der durch Gl. (2.17) gegebenen Grenzen bleiben müssen

$$x_i = x_{iu} + (x_{io} - x_{iu}) \sin^2 y_i . \quad (2.22)$$

Durch diese Transformationen wird erreicht, daß die Optimierung in einem "unbeschränkten y -Bildraum" durchgeführt werden kann. Der Term der die Ungleichheitsbedingungen in Gl. (2.20) berücksichtigt, kann daher weggelassen werden.

Wird außerdem $1/\sqrt{r_k}$ durch R_k ersetzt, dann ist

$$P(\underline{y}, R_k) = Z(\underline{y}) + R_k \sum_{j=1}^p h_j^2(\underline{y}) . \quad (2.23)$$

Nach Abschluß der Optimierung im Bildraum erhält man durch Rücktransformation den gesuchten optimalen Parametervektor $\underline{x}^*$.

Für die Gleichheitsbedingung $h_j(\underline{x})$ ($j=1,2,\dots,p$) ergeben sich aus den Gln. (2.19) und (2.15)

$$h_j(\underline{x}) = a_j(\underline{x}) - c_j = 0 \quad (j=1,2,\dots,p) \quad (2.24)$$

oder durch Einführen der Transformationen aus Gl.(2.21) bzw. Gl. (2.22)

$$h_j(\underline{y}) = a_j(\underline{y}) - c_j = 0 \quad (j=1,\dots,p) . \quad (2.25)$$

3.3 Der Gewichtungsfaktor R_k

Die Gewichtungsfaktoren R_k in Gl. (2.23) bilden im Gegensatz zu r_k eine monoton zunehmende Folge von positiven Werten. Auf diese Weise werden die Nebenbedingungen immer stärker berücksichtigt. Es taucht hier die Frage auf, welchen Anfangswert die Folge haben soll und welche Gesetzmäßigkeit zwischen aufeinanderfolgenden Werten existieren soll.

Bei der Wahl des Startvektors R_1 ist folgendes zu beachten:

Wird R_1 zu klein gewählt, überwiegt in der Gl. (2.23) die ursprüngliche Zielfunktion $Z(\underline{y})$ und die Nebenbedingungen werden sehr stark verletzt. Wird R_1 zu groß gewählt, überwiegt der zweite Term auf der rechten Seite der Gl. (2.23) und die Optimierung konzentriert sich hauptsächlich auf die Minimierung der Straffunktion. In diesem zweiten Fall entstehen längs der durch die Nebenbedingungen gegebenen Trajektorien im Parameterraum sehr enge Schluchten, wodurch die modifizierte Zielfunktion nur noch eine sehr geringe Ähnlichkeit mit der ursprünglichen hat. In beiden Fällen wird das Minimum des ersten Teilproblems weit vom gesuchten Minimum entfernt sein und daher zu einer schlechten Konvergenz führen.

Da eine notwendige Bedingung für die Minimierung von $P(\underline{y},R)$ das Verschwinden der Ableitungen 1. Ordnung ist, wird in [28] vorge-

schlagen, R_1 so zu wählen, daß der Betrag des Gradienten von $P(\underline{y}, R)$ im Startpunkt $\underline{y}_0$ minimal wird. Dadurch erreicht man, daß das Minimum des ersten Teilproblems nahe beim Startpunkt $\underline{y}_0$ liegt. R_1 muß also die Bedingung

$$|\nabla P(\underline{y}_0, R_1)| = \min_{R=0} |\nabla P(\underline{y}_0, R)| \quad (2.26)$$

erfüllen. Der Betrag des Gradienten ist

$$|\nabla P(\underline{y}, R)| = \sqrt{\sum_{i=1}^n \left(\frac{\partial P}{\partial y_i} \right)^2} \quad (2.27)$$

Hierin ist, mit $P(\underline{y}, R)$ aus Gl. (2.23),

$$\frac{\partial P}{\partial y_i} = \frac{\partial Z}{\partial y_i} + 2R \sum_{j=1}^p h_j \frac{\partial h_j}{\partial y_i} \quad (2.28)$$

Das Minimum von $|\nabla P(\underline{y}, R)|$ erhält man für

$$\frac{\partial}{\partial R} |\nabla P(\underline{y}, R)| = \frac{1}{|\nabla P(\underline{y}, R)|} \cdot \sum_{i=1}^n \frac{\partial P}{\partial y_i} \frac{\partial^2 P}{\partial y_i \partial R} = 0 \quad (2.29)$$

Hierin ist, durch Ableitung der Gl. (2.28) nach R ,

$$\frac{\partial^2 P}{\partial y_i \partial R} = 2 \sum_{j=1}^p h_j \frac{\partial h_j}{\partial y_i} \quad (2.30)$$

Aus den Gln. (2.28), (2.29) und (2.30) erhält man als Bedingung für R_1

$$\sum_{i=1}^n \left\{ \left(\frac{\partial Z}{\partial y_i} + 2 R_1 \sum_{j=1}^p h_j \frac{\partial h_j}{\partial y_i} \right) \sum_{j=1}^p h_j \frac{\partial h_j}{\partial y_i} \right\} = 0 \quad ,$$

woraus schließlich

$$R_1 = - \frac{\sum_{i=1}^n \left(\frac{\partial Z}{\partial y_i} \sum_{j=1}^p h_j \frac{\partial h_j}{\partial y_i} \right)}{2 \sum_{i=1}^n \left(\sum_{j=1}^p h_j \frac{\partial h_j}{\partial y_i} \right)^2} \quad (2.31)$$

folgt.

Das o.g. Kriterium für die Ermittlung des Anfangswertes der Folge von Gewichtungsfaktoren kann natürlich nur dann verwendet wer-

den, wenn die Gl. (2.31) einen Wert $R_1 > 0$ liefert. Für den Fall, daß $R_1 \leq 0$ ist, hat sich auf Grund der berechneten Beispielen erwiesen, daß eine günstige Wahl im Bereich $10 \leq R_1 \leq 100$ liegt.

Die Folge von Gewichtungsfaktoren wird, ausgehend von R_1 , durch einen konstanten Faktor c erzeugt, d.h.

$$R_{k+1} = c R_k. \quad (2.32)$$

Die Wahl von c ist nicht ausschlaggebend für die Konvergenz der Methode. Der Grund dafür ist der, daß für ein großes c mehr Iterationen je Teilproblem benötigt werden, für ein kleines c aber mehr Teilprobleme gelöst werden müssen. Die Gesamtzahl der Iterationen, also (Iterationen je Teilproblem) $\cdot$ (Anzahl der Teilprobleme), ist nahezu konstant und unabhängig von c . Es wird hier $c = 10$ verwendet.

Die Anzahl der Optimierungszyklen mit einem jeweils größeren R_k gemäß Gl. (2.32) wird so lange fortgesetzt bis das Abbruchkriterium

$$\sum_{j=1}^p \frac{h_j(\underline{x})}{c_j} < \epsilon \quad (2.33)$$

erfüllt ist. ϵ ist dabei eine Schranke, mit der die gewünschte Genauigkeit des Koeffizientenvergleichs beliebig vorgegeben werden kann.

3.4 Lösung des unbeschränkten Optimierungsproblems

In den vorangegangenen Abschnitten wurde gezeigt, wie das Optimierungsproblem mit Nebenbedingungen in eine Folge von unbeschränkten Optimierungen transformiert werden kann. Für die Lösung eines unbeschränkten Optimierungsproblems sind mehrere Verfahren und die zugehörigen Programme entwickelt worden. Eine gute Übersicht und ein Leistungsvergleich der verschiedenen Verfahren wird in [29] gegeben.

Die Optimierungsverfahren lassen sich in solche für allgemeine Zielfunktionen und solche, bei denen die Zielfunktion die Form einer Quadratsumme besitzt, aufteilen. Außerdem kann in jeder dieser Gruppen zwischen Verfahren, die nur die Zielfunktion benötigen und Verfahren, die außer der Zielfunktion auch den Gradienten der Zielfunktion benötigen, unterschieden werden.

Die modifizierte Zielfunktion Gl. (2.23) läßt sich, wie noch gezeigt werden soll, in die Form einer Quadratsumme bringen. Es können hier also außer den allgemeinen Verfahren auch die speziellen Verfahren für die Minimierung von Quadratsummen verwendet werden. Da die Ableitungen der Zielfunktion, mit der im Kapitel I, Abschnitt 3 beschriebenen Methode sehr leicht erzeugt werden können, ist es möglich, auch solche Verfahren einzusetzen, die den Gradienten der Zielfunktion benötigen. Für die Auswahl eines bestimmten Optimierungsverfahrens gibt es daher keinerlei Beschränkungen.

Im folgenden soll gezeigt werden, wie die Empfindlichkeitsoptimierung mit Hilfe eines allgemeinen Verfahrens und mit Hilfe eines Verfahrens für die Minimierung von Quadratsummen gelöst werden kann. Da es bei der Optimierung im allgemeinen von Vorteil ist, möglichst viel Information über die Zielfunktion zu besitzen, werden in beiden Fällen nur solche Verfahren berücksichtigt, die die Ableitungen der Zielfunktion benötigen.

Als Zielfunktion $Z(\underline{y})$ in Gl. (2.23) dient das Empfindlichkeitsmaß Gl. (2.7). Hierin ist zunächst die Variablentransformation gemäß Gl. (2.21) bzw. Gl. (2.22) einzuführen

$$M(\underline{y}) = \int_{\omega_a}^{\omega_b} \sum_{\nu=1}^n |S_{x_\nu}^F(\underline{y}, j\omega)|^2 . \quad (2.34)$$

Das Integral in Gl. (2.34) kann mit numerischen Integrationsmethoden (z.B. nach Simpson) ausgewertet werden. Der Einfachheit halber soll jedoch hier das Integral lediglich in eine endliche Summe umgewandelt werden. Im Frequenzintervall (ω_a, ω_b) werden dazu die Frequenzstützpunkte ω_l für $l = 1, 2, \dots, q$ eingeführt. Damit ist

$$M(\underline{y}) = \sum_{l=1}^q \sum_{\mu=1}^n |S_{x_{\mu}}^F(\underline{y}, \omega_l)|^2. \quad (2.35)$$

Es ist darauf zu achten, daß in den Gln. (2.34) und (2.35) die Empfindlichkeiten der Netzwerkfunktion F bezüglich der eigentlichen Netzwerkparameter x und nicht bezüglich der transformierten Parameter y einzusetzen ist. Es ist nämlich die Empfindlichkeit bezüglich der tatsächlichen Netzwerkparameter und nicht bezüglich irgendwelchen Transformationen davon zu minimieren.

Wird die Gl. (2.35) als Zielfunktion $Z(\underline{y})$ in Gl. (2.23) eingeführt, ist die modifizierte Zielfunktion

$$P(\underline{y}, R_k) = \sum_{l=1}^q \sum_{\mu=1}^n |S_{x_{\mu}}^F(\underline{y}, \omega_l)|^2 + R_k \sum_{j=1}^p h_j^2(\underline{y}). \quad (2.36)$$

Im Abschnitt 2.3 dieses Kapitels wurden Überlegungen bezüglich der Wahl der Netzwerkfunktion angestellt. Der wohl häufigste Fall in der Netzwerksynthese ist die Approximation einer Dämpfungsvorschrift. Aus diesem Grunde soll im folgenden die Dämpfung

$$a(\underline{x}, \omega) = -\ln |H(\underline{x}, j\omega)| \quad (2.37)$$

in Neper, an Stelle der Netzwerkfunktion F in Gl. (2.36) verwendet werden. Da die Dämpfung eine reelle Funktion ist, können die Betragsstriche in Gl. (2.36) weggelassen werden. Damit gilt

$$P(\underline{y}, R_k) = \sum_{l=1}^q \sum_{\mu=1}^n (S_{x_{\mu}}^a(\underline{y}, \omega_l))^2 + R_k \sum_{j=1}^p h_j^2(\underline{y}). \quad (2.38)$$

3.4.1 Lösung mit Hilfe eines allgemeinen Optimierungsverfahrens

Außer der Zielfunktion gemäß Gl. (2.38) wird für ein allgemeines Optimierungsverfahren auch der Gradient der Zielfunktion

$$\nabla P(\underline{y}) = \left(\frac{\partial P}{\partial y_1}, \frac{\partial P}{\partial y_2}, \dots, \frac{\partial P}{\partial y_n} \right)^t \quad (2.39)$$

benötigt. Es sind also die Ableitungen der Zielfunktion bezüg-

lich aller transformierten Parameter zu berechnen. Da die Empfindlichkeiten und die Nebenbedingungen zunächst als Funktionen der Netzwerkparameter x_i ($i=1,2,\dots,n$) bekannt sind, berechnet man die Ableitungen von P bezüglich y_i ($i=1,2,\dots,n$) mit Hilfe der Kettenregel

$$\frac{\partial P}{\partial y_v} = \frac{\partial P}{\partial x_v} \frac{\partial x_v}{\partial y_v} \quad (2.40)$$

Hierin ist

$$\frac{\partial P}{\partial x_v} = 2 \sum_{l=1}^q \sum_{\mu=1}^n S'_{x_\mu}{}^a(\underline{x}, \omega_l) \frac{\partial}{\partial x_v} S'_{x_\mu}{}^a(\underline{x}, \omega_l) + 2R_{k,j=1}^p h_j(\underline{x}) \frac{\partial h_j(\underline{x})}{\partial x_v} \quad (2.41)$$

Außer den Empfindlichkeiten sind demnach auch die Ableitungen der Empfindlichkeiten und die Ableitungen der Nebenbedingungen in einem beliebigen Punkt des Parameterraums zu berechnen. Im Abschnitt 4 wird gezeigt, wie diese Werte berechnet werden können.

3.4.2 Lösung mit Hilfe eines Optimierungsverfahrens für Quadratsummen

Für den Sonderfall, daß die Zielfunktion eine Quadratsumme nichtlinearer Funktionen ist, sind spezielle Optimierungsverfahren entwickelt worden [26], [29].

Die Quadratsumme in Gl. (2.38) kann auch als Skalarprodukt

$$P(\underline{y}) = \underline{f}(\underline{y}) \cdot \underline{f}(\underline{y}) \quad (2.42)$$

des Vektors

$$\underline{f}(\underline{y}) = (f_1(\underline{y}), f_2(\underline{y}), \dots, f_m(\underline{y}))$$

mit sich selbst interpretiert werden, wenn

$$\underline{f}(\underline{y}) = (S'_{x_1}{}^{a_1}, S'_{x_2}{}^{a_1}, \dots, S'_{x_n}{}^{a_1}, S'_{x_1}{}^{a_2}, \dots, S'_{x_n}{}^{a_2}, \dots, S'_{x_1}{}^{a_q}, \dots, S'_{x_n}{}^{a_q}, \sqrt{R_k} h_1, \sqrt{R_k} h_2, \dots, \sqrt{R_k} h_p) \quad (2.43)$$

ist. Hierin ist $S'_{x_\mu}^{a_1} = S'_{x_\mu}^a(\underline{y}, \omega_1)$ die halbrerelative Empfindlichkeit der Dämpfung im Frequenzpunkt ω_1 in Bezug auf den Netzwerkparameter x_μ und $h_1, h_2, \dots, h_p$ sind die Nebenbedingungen nach Gl. (2.25). Die Dimension des Vektors $\underline{f}(\underline{y})$ ist

$$m = n \cdot q + p. \quad (2.44)$$

Außer dem Vektor $\underline{f}(\underline{y})$ wird die Jacobische Matrix

$$\underline{J}(\underline{y}) = \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} & \dots & \frac{\partial f_1}{\partial y_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1} & \frac{\partial f_m}{\partial y_2} & \dots & \frac{\partial f_m}{\partial y_n} \end{bmatrix} \quad (2.45)$$

benötigt. Die einzelnen Elemente dieser Matrix erhält man wiederum mit Hilfe der Kettenregel:

$$\frac{\partial f_i}{\partial y_v} = \frac{\partial f_i}{\partial x_v} \frac{\partial x_v}{\partial y_v}, \quad (2.46)$$

wobei

$$\frac{\partial f_i}{\partial x_v} = \begin{cases} \frac{\partial}{\partial x_v} S'_{x_1}^a(\underline{x}, \omega_1) & \text{für } v = 1, 2, \dots, n \\ & \text{und } i = 1, 2, \dots, n \\ \frac{\partial}{\partial x_v} S'_{x_{i-n}}^a(\underline{x}, \omega_2) & \text{für } v = 1, 2, \dots, n \\ & \text{und } i = n+1, n+2, \dots, 2n \\ \vdots & \\ \frac{\partial}{\partial x_v} S'_{x_{i-(q-1)n}}^a(\underline{x}, \omega_q) & \text{f. } v = 1, 2, \dots, n \\ & \text{und } i = (q-1)n+1, \dots, qn \\ \sqrt{R_k} \frac{\partial}{\partial x_v} h_{i-qn} & \text{für } v = 1, 2, \dots, n \\ & \text{und } i = qn+1, \dots, qn+p \end{cases} \quad (2.47)$$

gilt und $\partial x_v / \partial y_v$ aus der Transformationsgleichung (2.21) oder (2.22) folgt. Wie in 3.4.1 werden auch hier außer den Empfindlichkeiten selbst zusätzlich deren Ableitungen und die Ableitungen der Nebenbedingungen benötigt.

4. Die Berechnung der Empfindlichkeiten und deren Ableitungen

In diesem Abschnitt soll gezeigt werden, wie die in der Zielfunktion Gl. (2.38) und im Vektor Gl. (2.43) enthaltenen Empfindlichkeiten berechnet werden können. Es soll außerdem gezeigt werden, wie die Ableitungen der Empfindlichkeiten für die Elemente des Gradienten, Gl. (2.41), und für die Jacobische Matrix, Gl. (2.45) berechnet werden können.

Während der Optimierung müssen sowohl die Zielfunktion als auch die Ableitungen für einen beliebigen Wert des Parametervektors $\underline{x}$ bzw. des transformierten Parametervektors $\underline{y}$ berechnet werden. Wie bereits im Kapitel I begründet, geschieht das am einfachsten auf Grund einer Netzwerkfunktion, die in symbolischer Form gegeben ist.

4.1 Berechnung der Empfindlichkeiten

Die halbrelative Empfindlichkeit der Dämpfung Gl. (2.37) läßt sich aus dem Realteil der vollrelativen Empfindlichkeit der Übertragungsfunktion berechnen. Gemäß Gl. (1.7) ist

$$S'_{x_{\mu}}^a(\underline{x}, \omega) = -\operatorname{Re} \left\{ S''_{x_{\mu}}^H(\underline{x}, j\omega) \right\}. \quad (2.48)$$

Die vollrelative Empfindlichkeit der Übertragungsfunktion kann nun an Hand der Gl. (1.67) auf sehr einfache Weise aus der Übertragungsfunktion in symbolischer Form gewonnen werden.

4.2 Berechnung der Ableitungen der Empfindlichkeiten

In den Gln. (2.41) und (2.47) werden die Ableitungen der Empfindlichkeiten benötigt. Aus Gl. (2.48) folgt

$$\frac{\partial}{\partial x_{\nu}} S'_{x_{\mu}}^a = -\operatorname{Re} \left\{ \frac{\partial}{\partial x_{\nu}} S''_{x_{\mu}}^H \right\}. \quad (2.49)$$

Da es sich beim Ausdruck in den rechteckigen Klammern nicht um die eigentliche Empfindlichkeit zweiter Ordnung, sondern um die Ableitung der vollrelativen Empfindlichkeit erster Ordnung handelt, kann hier nicht direkt die Gl. (1.71) eingesetzt werden. Die Ableitung der vollrelativen Empfindlichkeit kann aber, ähnlich wie die Empfindlichkeit zweiter Ordnung, aus den Empfindlichkeiten erster und zweiter Ordnung von Zähler und Nenner der Übertragungsfunktion

$$H(\underline{x}, p) = \frac{Z(\underline{x}, p)}{N(\underline{x}, p)} \quad (2.50)$$

berechnet werden.

Nach Gl. (1.9) gilt

$$S_{x_{\mu}}^H = \frac{x_{\mu}}{Z} \frac{\partial Z}{\partial x_{\mu}} - \frac{x_{\mu}}{N} \frac{\partial N}{\partial x_{\mu}} \quad (2.51)$$

Daher ist

$$\begin{aligned} \frac{\partial}{\partial x_{\nu}} S_{x_{\mu}}^H &= \left(\frac{1}{Z} \frac{\partial x_{\mu}}{\partial x_{\nu}} - \frac{x_{\mu}}{Z^2} \frac{\partial Z}{\partial x_{\nu}} \right) \frac{\partial Z}{\partial x_{\mu}} + \frac{x_{\mu}}{Z} \frac{\partial^2 Z}{\partial x_{\mu} \partial x_{\nu}} \\ &- \left(\frac{1}{N} \frac{\partial x_{\mu}}{\partial x_{\nu}} - \frac{x_{\mu}}{N^2} \frac{\partial N}{\partial x_{\nu}} \right) \frac{\partial N}{\partial x_{\mu}} - \frac{x_{\mu}}{N} \frac{\partial^2 N}{\partial x_{\mu} \partial x_{\nu}} \quad (2.52) \end{aligned}$$

Es ist nun zweckmäßig, die folgenden zwei Fälle zu unterscheiden:

Fall 1 : $\nu = \mu$

Für $\nu = \mu$ ist $\partial x_{\mu} / \partial x_{\nu} = 1$ ($\mu = \nu = 1, 2, \dots, n$). Da im Zähler und Nenner die Parameter jeweils nur in ihrer ersten Potenz enthalten sind, gilt außerdem:

$$\frac{\partial^2 Z}{\partial x_{\mu} \partial x_{\nu}} = \frac{\partial^2 Z}{\partial x_{\mu}^2} \equiv 0$$

und

$$\frac{\partial^2 N}{\partial x_{\mu} \partial x_{\nu}} = \frac{\partial^2 N}{\partial x_{\mu}^2} \equiv 0 \quad .$$

Damit wird aus Gl. (2.52)

$$\frac{\partial}{\partial x_{\mu}} S^H_{x_{\mu}} = \frac{1}{Z} \frac{\partial Z}{\partial x_{\mu}} - \frac{1}{N} \frac{\partial N}{\partial x_{\mu}} - \frac{x_{\mu}}{Z^2} \left(\frac{\partial Z}{\partial x_{\mu}} \right)^2 + \frac{x_{\mu}}{N^2} \left(\frac{\partial N}{\partial x_{\mu}} \right)^2 \quad (2.53)$$

und auf Grund von Gl. (2.51)

$$\frac{\partial}{\partial x_{\mu}} S^H_{x_{\mu}} = \frac{1}{x_{\mu}} \left\{ S^H_{x_{\mu}} - (S^Z_{x_{\mu}})^2 + (S^N_{x_{\mu}})^2 \right\} \quad (2.54)$$

Fall 2 : $v \neq \mu$

Da bei der Berechnung des Gradienten, um den es hier geht, die Parameter $x_1, x_2, \dots, x_n$ unabhängige Variablen darstellen, ist $\partial x_{\mu} / \partial x_v = 0$ ($\mu, v=1, 2, \dots, n$). Damit wird aus Gl. (2.52)

$$\frac{\partial}{\partial x_v} S^H_{x_{\mu}} = - \frac{x_{\mu}}{Z^2} \frac{\partial Z}{\partial x_v} \frac{\partial Z}{\partial x_{\mu}} + \frac{x_{\mu}}{N^2} \frac{\partial N}{\partial x_v} \frac{\partial N}{\partial x_{\mu}} + \frac{x_{\mu}}{Z} \frac{\partial^2 Z}{\partial x_{\mu} \partial x_v} - \frac{x_{\mu}}{N} \frac{\partial^2 N}{\partial x_{\mu} \partial x_v} \quad (2.55)$$

oder

$$\frac{\partial}{\partial x_v} S^H_{x_{\mu}} = - \frac{1}{x_v} (S^Z_{x_v} S^Z_{x_{\mu}} - S^N_{x_v} S^N_{x_{\mu}}) + \frac{x_{\mu}}{Z} S^Z_{x_{\mu} x_v} - \frac{x_{\mu}}{N} S^N_{x_{\mu} x_v} \quad (2.56)$$

Die Ableitungen der vollrelativen Empfindlichkeiten können also auf Grund der Gln. (2.54) und (2.56) mit Hilfe der vollrelativen Empfindlichkeit der Übertragungsfunktion gemäß Gl. (1.67), der vollrelativen Zähler- und Nennerempfindlichkeit gemäß Gln. (1.63) und (1.64) und der Empfindlichkeit zweiter Ordnung von Zähler und Nenner nach Gln. (1.72) und (1.73) berechnet werden.

Wegen Gl. (1.77) ist, wie aus Gl. (2.56) sofort erkannt werden kann,

$$\frac{\partial}{\partial x_{\mu}} S^H_{x_v} = \frac{x_v}{x_{\mu}} \frac{\partial}{\partial x_v} S^H_{x_{\mu}} \quad (2.57)$$

Es genügt daher, wenn die Ableitungen in Gl. (2.56) für

$\mu = 1, 2, \dots, n$ und $v > \mu$ berechnet werden. Für $v < \mu$ erhält man sie dann gemäß Gl. (2.57).

4.3 Bemerkung

Im Empfindlichkeitsmaß nach Gl. (2.7) wurde die halbrelative Empfindlichkeit verwendet. Selbstverständlich könnte an ihrer Stelle auch die vollrelative Empfindlichkeit eingesetzt werden. Daß hier jedoch die halbrelative Empfindlichkeit verwendet wurde, liegt lediglich darin, daß man mit ihnen zu einfacheren Ausdrücken der Ableitungen Gln. (2.54) und (2.56) kommt.

5. Die Freiheitsgrade bei der Empfindlichkeitsoptimierung

Am Anfang des Kapitels wurde darauf hingewiesen, daß eine Optimierung der Empfindlichkeit nur dann möglich ist, wenn bei der Wahl der Netzwerkparameter Freiheitsgrade vorhanden sind. Im folgenden soll gezeigt werden, daß dies nicht immer der Fall ist.

Wenn Zähler und Nennerterme der Netzwerkfunktion Gl. (1.54) nach den p -Potenzen geordnet werden, gilt

$$F(\underline{x}, p) = \frac{Z(\underline{x}, p)}{N(\underline{x}, p)} = \frac{\sum_{i=0}^l \left\{ \sum_{j=1}^{r_i} a_{ij}(\underline{x}) \right\} p^i}{\sum_{i=0}^m \left\{ \sum_{j=1}^{s_i} b_{ij}(\underline{x}) \right\} p^i} \quad (2.58)$$

Hierin ist l der Zähler- und m der Nennergrad, r_i die Anzahl der Terme mit i -ter Potenz von p im Zähler und s_i die Anzahl der Terme mit i -ter Potenz von p im Nenner.

Die vollrelative Empfindlichkeit der Netzwerkfunktion $F(\underline{x}, p)$ bezüglich eines Netzwerkparameters x_μ ($\mu=1, 2, \dots, n$) kann nun, ähnlich wie in Kap. I, Abschnitt 3.2 gezeigt wurde, aus den Termen von Zähler und Nenner berechnet werden, die x_μ als Faktor enthalten

$$S_{x_\mu}^F(\underline{x}, p) = \frac{1}{Z(\underline{x}, p)} \sum_{i=0}^l \left\{ \sum_{j \in K_{\mu i}} a_{ij}(\underline{x}) \right\} p^i - \frac{1}{N(\underline{x}, p)} \sum_{i=0}^m \left\{ \sum_{j \in L_{\mu i}} b_{ij}(\underline{x}) \right\} p^i \quad (2.59)$$

Hierin ist

$$K_{\mu i} = \{j \mid a_{ij}(\underline{x}) = f(x_\mu)\} \quad (i=0, 1, \dots, l) \quad (2.60)$$

die Menge aller j -Indizes der Zählerterme, und

$$L_{\mu i} = \{j \mid b_{ij}(\underline{x}) = f(x_\mu)\} \quad (i=0, 1, \dots, m) \quad (2.61)$$

die Menge aller j -Indizes der Nennerterme mit der i -ten Potenz

in p , die x_u als Faktor enthalten.

Gl. (2.59) unterscheidet sich von Gl. (1.67) lediglich darin, daß in ihr die Zähler- und Nennerterme nach den p -Potenzen geordnet wurden.

Da sich die Netzwerkfunktion, und damit auch der Zähler und Nenner, während der Optimierung nicht ändern dürfen, kann die vollrelative Empfindlichkeit aus Gl. (2.59) allein als Funktion der a_{ij} 's, der b_{ij} 's und des Frequenzparameters p angesehen werden. Das gleiche gilt dann auch, wie durch Einsetzen der Gl. (2.59) in Gl. (2.7) sofort erkannt werden kann, für das integrale Empfindlichkeitsmaß, d.h.

$$M(\underline{x}) = f(\text{alle } a_{ij} \text{ und alle } b_{ij}) . \quad (2.62)$$

Die Zähler- und Nennerterme sind aber Einschränkungen auf Grund des Koeffizientenvergleichs unterworfen. Sind diese Einschränkungen so, daß jedes a_{ij} und jedes b_{ij} einen festen Wert besitzt, dann sind die Freiheitsgrade durch die Nebenbedingungen ausgeschöpft. Das Empfindlichkeitsmaß hat dann einen Wert, der nur von der zu erfüllenden Vorschrift und nicht von den Netzwerkparametern abhängt. Dieser Fall tritt mit Sicherheit dann auf, wenn die Koeffizienten von p im Zähler und Nenner der Netzwerkfunktion aus jeweils nur einem Term bestehen. Diese Koeffizienten müssen auf Grund der Nebenbedingungen (Koeffizientenvergleich), nämlich einen vorgegebenen Wert beibehalten. Diese Tatsache kann am einfachsten im folgenden Beispiel verdeutlicht werden.

Beispiel 2.1: Bild 2.1 zeigt ein aktives Netzwerk nach Åkerberg und Mossberg zur Realisierung einer Übertragungsfunktion zweiten Grades [30].

Um zu erreichen, daß der Koeffizient der höchsten p -Potenz im Nenner der Übertragungsfunktion unabhängig von den Parameterwerten und gleich eins wird, werden hier bewußt nicht alle Elemente einheitlich, z.B. durch Widerstand und Kapazität, beschrieben.

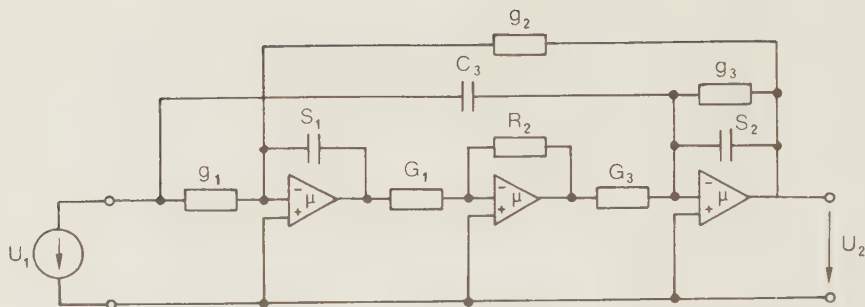


Bild 2.1: Aktives Netzwerk nach Åkerberg und Mossberg

Werden ideale Operationsverstärker (Verstärkung $\mu \rightarrow \infty$) vorausgesetzt, ist die Übertragungsfunktion

$$H(p) = \frac{U_2}{U_1} = \frac{Z}{N} = \frac{-S_2 C_3 p^2 - g_1 G_1 R_2 G_3 S_1 S_2}{p^2 + g_3 S_2 p + g_2 G_1 R_2 G_3 S_1 S_2} . \quad (2.63)$$

Mit den Bezeichnungen aus Gl. (2.58) sind die Zähler- und Nennerkoeffizienten:

$$\begin{aligned} a_{01} &= -g_1 G_1 R_2 G_3 S_1 S_2 , & b_{01} &= g_2 G_1 R_2 G_3 S_1 S_2 , \\ a_{11} &= 0 , & b_{11} &= g_3 S_2 , \\ a_{21} &= -S_2 C_3 , & b_{21} &= 1 . \end{aligned} \quad (2.64)$$

Diese Schaltung soll nun in Bezug auf ihre Gesamtempfindlichkeit optimiert werden. Es ist dabei die Übertragungsfunktion mit numerisch vorgegebenen Koeffizienten

$$H(p) = \frac{a'_2 p^2 + a'_0}{p^2 + b'_1 p + b'_0} \quad (2.65)$$

zu realisieren.

Mit Hilfe der Gl. (2.59) können die vollrelativen Empfindlichkeiten der Übertragungsfunktion bezüglich aller Parameter berechnet werden. Es ergeben sich die Beziehungen

$$S''_{g_1}^H = \frac{1}{Z} a_{01} , \quad S''_{g_2}^H = -\frac{1}{N} b_{01} ,$$

$$S''_{g_3}^H = -\frac{1}{N} b_{11}p , \quad S''_{G_1}^H = \frac{1}{Z} a_{01} - \frac{1}{N} b_{01} ,$$

$$S''_{R_2}^H = \frac{1}{Z} a_{01} - \frac{1}{N} b_{01} , \quad S''_{G_3}^H = \frac{1}{Z} a_{01} - \frac{1}{N} b_{01} ,$$

$$S''_{S_1}^H = \frac{1}{Z} a_{01} - \frac{1}{N} b_{01} , \quad S''_{S_2}^H = \frac{1}{Z} (a_{21}p^2 + a_{01}) - \frac{1}{N} (b_{11}p + b_{01}) ,$$

$$S''_{C_3}^H = \frac{1}{Z} a_{21} p^2 .$$

Die Einzelempfindlichkeiten an einem Frequenzpunkt und mit ihnen auch das Empfindlichkeitsmaß, hängen also nur von den Koeffizienten aus Gl. (2.64) ab.

Da gleichzeitig die vorgeschriebene Übertragungsfunktion Gl. (2.65) eingehalten werden muß, gelten folgende Nebenbedingungen:

$$a_{01} = a'_0 ,$$

$$a_{21} = a'_2 ,$$

$$b_{01} = b'_0 ,$$

$$b_{11} = b'_1 .$$

Jeder einzelne Koeffizient der symbolischen Übertragungsfunktion Gl. (2.63) hat damit einen Wert, der durch die zu realisierende Übertragungsfunktion Gl. (2.65) fest vorgegeben ist. Wie in Gl. (2.63) zu erkennen ist, können zwar die Bauelemente G_1 , R_2 , G_3 und S_1 mit der Bedingung, daß deren Produkt konstant bleibt, frei gewählt werden, aber die Empfindlichkeiten können damit nicht beeinflusst werden. Eine Empfindlichkeitsoptimierung ist daher in diesem Fall nicht möglich.

Auch dann, wenn die Zähler- und Nennerkoeffizienten von p nicht nur aus jeweils einem Term bestehen, sondern sich aus der Summe

mehrerer Terme zusammensetzen, ist eine Optimierung der Gesamtempfindlichkeit nicht immer möglich. Dies soll an Hand des folgenden Beispiels gezeigt werden.

Beispiel 2.2:

Der aktive Tiefpaß zweiten Grades aus Bild 1.7 soll in Bezug auf die Gesamtempfindlichkeit optimiert werden. Es ist dabei die Übertragungsfunktion mit numerisch vorgegebenen Koeffizienten

$$H(p) = \frac{a'_0}{p^2 + b'_1 p + b'_0} \quad (2.66)$$

einzuhalten.

Wird $S_1 = 1/C_1$ und $S_2 = 1/C_2$ in Gl. (1.56) eingeführt, erhält man die Übertragungsfunktion

$$H(p) = \frac{Z}{N} = \frac{g_1 g_2 S_1 S_2 k_1 k_2}{p^2 + (g_1 S_1 + g_2 S_2 - g_2 S_2 k_1 k_2) p + g_1 g_2 S_1 S_2} \quad (2.67)$$

oder mit den Bezeichnungen aus Gl. (2.58)

$$H(p) = \frac{Z}{N} = \frac{a_{01}}{p^2 + (b_{11} + b_{12} + b_{13}) p + b_{01}} \quad (2.68)$$

Hierin ist

$$\begin{aligned} a_{01} &= g_1 g_2 S_1 S_2 k_1 k_2 , \\ b_{01} &= g_1 g_2 S_1 S_2 , \\ b_{11} &= g_1 S_1 , \\ b_{12} &= g_1 S_2 , \\ b_{13} &= -g_2 S_2 k_1 k_2 , \\ b_{21} &= 1 . \end{aligned} \quad (2.69)$$

Das Empfindlichkeitsmaß kann gemäß Gl. (2.62) allein als Funktion der Koeffizienten aus Gl. (2.69) angesehen werden. Es erhebt sich nun die Frage, ob zumindest einer dieser Koeffizienten frei gewählt werden kann. Ist dies nicht der Fall, wird eine Optimierung nicht möglich sein. Zur Beantwortung dieser Frage sind zunächst die durch den Koeffizientenvergleich auferlegten Einschränkungen der Koeffizienten aus Gl. (2.69) zu betrachten. Diese sind

$$a_{01} = a'_0, \quad (2.70a)$$

$$b_{01} = b'_0, \quad (2.70b)$$

$$b_{11} + b_{12} + b_{13} = b'_1. \quad (2.70c)$$

Auf Grund dieser Gleichungen könnte der Schluß gezogen werden, daß eine Optimierung möglich ist, da die letzte von ihnen eine freie Wahl von b_{11} , b_{12} und b_{13} (mit der Einschränkung, daß deren Summe gleich b'_1 sein muß) zuläßt. Dies ist aber nicht so, da, wie aus den Gln. (2.69) ersichtlich, zusätzlich zu den Gln. (2.70a-c) noch folgende Bedingungen angegeben werden können:

$$b_{13} = -b_{12} \frac{a_{01}}{b_{01}} \quad (2.71a)$$

$$b_{11} b_{12} = b_{01} \quad (2.71b)$$

Die Gln. (2.70a-c) und (2.71a,b) bilden zusammen ein Gleichungssystem mit 5 unabhängigen Gleichungen und 5 Unbekannten. Die Koeffizienten a_{01} , b_{01} , b_{11} , b_{12} und b_{13} sind daher in Abhängigkeit von a'_0 , b'_0 und b'_1 eindeutig bestimmbar. Das Empfindlichkeitsmaß hat somit einen Wert, der nur noch von den Koeffizienten der vorgeschriebenen Übertragungsfunktion abhängt. Eine Empfindlichkeitsoptimierung der Schaltung ist also auch in diesem Fall nicht möglich.

Wird jedoch die in Gl. (2.70a) enthaltene Nebenbedingung aufgehoben, also auf die Einhaltung des konstanten Faktors im Zähler kein Wert gelegt, so ist in dem aus den Gln. (2.70 b,c) und (2.71a,b) gebildeten Gleichungssystem ein Freiheitsgrad vorhanden, der für die Empfindlichkeitsoptimierung ausgenutzt wer-

den kann (siehe 1. Beispiel im Abschnitt 6).

Im vorangehenden wurde gezeigt, daß die Optimierung der Empfindlichkeit einer Schaltung nicht immer möglich ist. Aus diesem Grund sollte, bevor mit einer Optimierung begonnen wird, erst einmal untersucht werden, ob zu den Nebenbedingungen aus dem Koeffizientenvergleich weitere Gleichungen für die Koeffizienten der symbolischen Netzwerkfunktion aufgestellt werden können. Nur dann, wenn das somit aufgestellte Gleichungssystem unbestimmt ist, kann eine Optimierung durchgeführt werden.

Das hier gesagte für die Optimierung der vollrelativen Empfindlichkeit einer komplexwertigen Netzwerkfunktion gilt natürlich, wie leicht einzusehen ist, in gleicher Weise für den Fall, daß an Stelle der vollrelativen, die halbrelative Empfindlichkeit, und/oder an Stelle einer komplexwertigen eine reellwertige Netzwerkfunktion betrachtet wird.

6. Praktische Beispiele

Das in den vorangegangenen Abschnitten beschriebene Verfahren für die Empfindlichkeitsoptimierung linearer Netzwerke wurde an einer Reihe von praktischen Beispielen erfolgreich erprobt. Die numerischen Rechnungen hierzu wurden an der Rechenanlage CDC 330 des Rechenzentrums der Universität Erlangen-Nürnberg durchgeführt. Für die Lösung des unbeschränkten Optimierungsproblems wurden wahlweise die am Lehrstuhl für Allgemeine und Theoretische Elektrotechnik entwickelten Programme FLETCHER-POWELL III und MODNEW (Gauß Newton) [29] verwendet.

1. Beispiel

Als erstes Beispiel soll der in Kapitel I, Abschnitt 3.1 und im Abschnitt 5 dieses Kapitels betrachtete aktive Tiefpaß zweiten Grades aus Bild 1.7 herangezogen werden. Diese Schaltung soll so dimensioniert werden, daß die Gesamttempfindlichkeit der Dämpfung in Neper ($a(\omega) = -\ln |H(j\omega)|$) im Frequenzbereich $\omega \in (0.5, 1.5)$ zu einem Minimum wird. Als Empfindlichkeitsmaß soll das Maß nach Gl. (2.7) mit der halbrelativen Empfindlichkeit der Dämpfung verwendet werden. Die Schaltung hat dabei die Butterworth-Charakteristik zweiter Ordnung.

$$H(p) = \frac{k}{p^2 + \sqrt{2}p + 1} \quad (2.72)$$

zu realisieren.

Im Abschnitt 5 dieses Kapitels wurde festgestellt, daß eine Empfindlichkeitsoptimierung dieser Schaltung nur dann möglich ist, wenn der konstante Faktor im Zähler der Übertragungsfunktion keinen vorgeschriebenen Wert einhalten muß. Aus diesem Grund soll hier vorausgesetzt werden, daß k in Gl. (2.72) einen beliebigen Wert annehmen darf. Wird der Parametervektor

$$\underline{x} = (g_1, g_2, s_1, s_2, k_1, k_2)^t \quad (2.73)$$

eingeführt, dann lautet die Übertragungsfunktion Gl. (2.67)

$$H(\underline{x}, p) = \frac{x_1 x_2 x_3 x_4 x_5 x_6}{p^2 + (x_1 x_3 + x_2 x_4 - x_2 x_4 x_5 x_6) p + x_1 x_2 x_3 x_4} \quad (2.74)$$

Der Koeffizientenvergleich der Übertragungsfunktionen Gl. (2.72) und (2.74) liefert die Nebenbedingungen

$$h_1(\underline{x}) = x_1 x_3 + x_2 x_4 - x_2 x_4 x_5 x_6 - \sqrt{2} = 0 \quad (2.75a)$$

und

$$h_2(\underline{x}) = x_1 x_2 x_3 x_4 - 1 = 0 \quad (2.75b)$$

Als Nebenbedingungen für die Parameterwerte soll hier lediglich deren Positivität gefordert werden, d.h. $x_1, x_2, \dots, x_6 \geq 0$. Es ist daher die Variablentransformation nach Gl. (2.21)

$$x_i = y_i^2 \quad (i = 1, 2, \dots, 6) \quad (2.76)$$

einzuführen. Die Optimierung findet dann im unbeschränkten y -Bildraum statt.

Die Dimensionierung der Schaltung vor der Optimierung sei durch folgenden Anfangswert des Parametervektors gegeben

$$\underline{x}_0 = \begin{bmatrix} 1,0496 \\ 0,5713 \\ 3,610 \\ 0,4619 \\ 2,0 \\ 5,0 \end{bmatrix}.$$

Im Frequenzbereich $\omega \in (0.5, 1.5)$ werden nun $q = 11$ gleichverteilte Frequenzstützpunkte eingeführt. Das Empfindlichkeitsmaß Gl. (2.35) hat damit vor der Optimierung den Wert $M(\underline{x}_0) = 306,9$.

Nach der Optimierung ergibt sich für das Empfindlichkeitsoptimale Netzwerk der Parametervektor

$$\underline{x}^* = \begin{bmatrix} 0,3205 \\ 1,0995 \\ 2,8622 \\ 0,9914 \\ 0,1778 \\ 3,0607 \end{bmatrix}.$$

und das Empfindlichkeitsmaß $M(\underline{x}^*) = 53,7$. Durch die Optimierung wurde also eine Verbesserung im Empfindlichkeitsverhalten des Netzwerks um den Faktor 5,72 erzielt.

Das Empfindlichkeitsverhalten des Netzwerks vor und nach der Optimierung ist aus dem Bild 2.2 ersichtlich. Hier wurde die Schoeffflersche Gesamtempfindlichkeit Gl. (2.3) der Dämpfung über den Frequenzbereich $\omega \in [0.5, 1.5]$ aufgetragen.

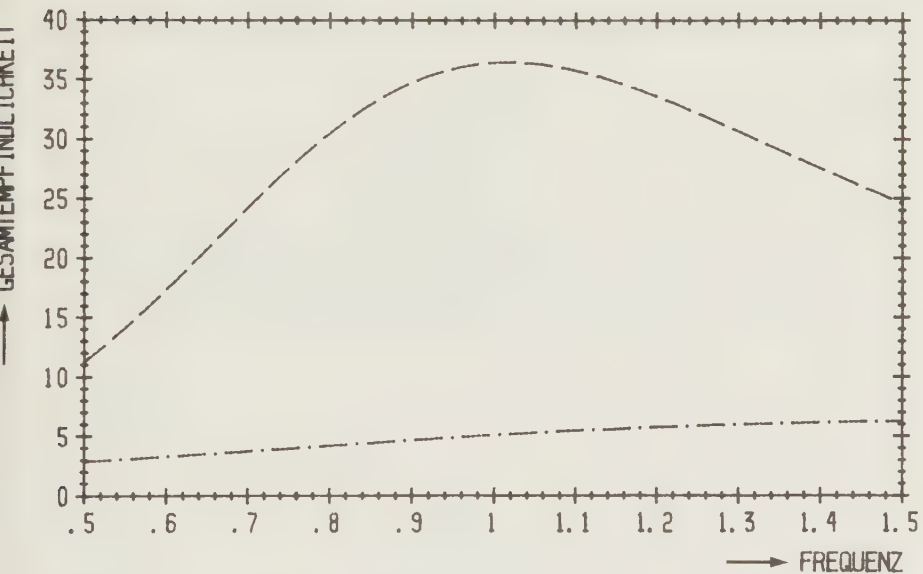


Bild 2.2: Gesamtempfindlichkeit des Netzwerks des ersten Beispiels

— — — : vor der Optimierung
— · — : nach der Optimierung

2. Beispiel

Als zweites Beispiel soll der bereits in [24] für die Empfindlichkeitsoptimierung herangezogene Sallen und Key-Tiefpaß betrachtet werden (Bild 2.3).

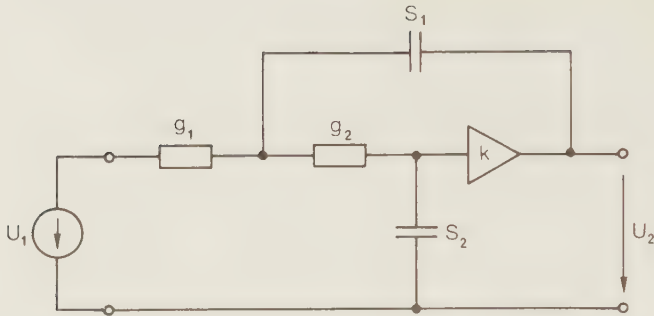


Bild 2.3: Sallen und Key-Tiefpaß

Der Tiefpaß soll so dimensioniert werden, daß die Gesamtempfindlichkeit der Dämpfung im Frequenzbereich $\omega \in (0.5, 1.5)$ zu einem Minimum wird. Als Empfindlichkeitsmaß soll das Maß nach Gl. (2.7) mit der halbrelativen Empfindlichkeit der Dämpfung verwendet werden. Die Schaltung in Bild 2.3 hat dabei die Butterworth-Charakteristik zweiter Ordnung

$$H(p) = \frac{10}{p^2 + \sqrt{2} p + 1} \quad (2.77)$$

zu realisieren.

Wird der Parametervektor

$$\underline{x} = (g_1, g_2, s_1, s_2, k)^t \quad (2.78)$$

eingeführt, liefert die Analyse der Schaltung die Übertragungsfunktion

$$H(\underline{x}, p) = \frac{x_1 x_2 x_3 x_4 x_5}{p^2 + (x_1 x_3 + x_2 x_3 + x_2 x_4 - x_2 x_4 x_5) + x_1 x_2 x_3 x_4} \quad (2.79)$$

Der Koeffizientenvergleich der Übertragungsfunktionen Gl. (2.77) und (2.79) liefert die Nebenbedingungen

$$h_1(\underline{x}) = x_1 x_2 x_3 x_4 x_5 - 10 = 0, \quad (2.80a)$$

$$h_2(\underline{x}) = x_1x_3 + x_2x_3 + x_2x_4 - x_2x_4x_5 - \sqrt{2} = 0, \quad (2.80b)$$

$$h_3(\underline{x}) = x_1x_2x_3x_4 - 1 = 0. \quad (2.80c)$$

Als Nebenbedingungen für die Parameterwerte soll hier, wie im 1. Beispiel, nur ihre Positivität gefordert werden. Wird die Variablentransformation $x_i = y_i^2$ ($i=1,2,\dots,5$) eingeführt, kann die Optimierung im unbeschränkten y -Bildraum stattfinden.

Die Dimensionierung der Schaltung vor der Optimierung sei durch folgenden Anfangswert des Parametervektors gegeben

$$\underline{x}_0 = \begin{bmatrix} 0,5542 \\ 1,0835 \\ 3,6102 \\ 0,4619 \\ 10,0 \end{bmatrix}.$$

Damit diese Schaltung mit der aus dem 1. Beispiel verglichen werden kann, werden im Frequenzbereich $\omega \in \{0.5, 1.5\}$ die gleichen elf Frequenzstützpunkte eingeführt. Das Empfindlichkeitsmaß Gl. (2.35) hat dann vor der Optimierung den Wert $M(\underline{x}_0) = 412,4$.

Nach der Optimierung ergibt sich für das empfindlichkeitsoptimale Netzwerk der Parametervektor

$$\underline{x}^* = \begin{bmatrix} 8,0691 \\ 0,2591 \\ 0,4606 \\ 1,0384 \\ 10,0 \end{bmatrix}$$

und das Empfindlichkeitsmaß $M(\underline{x}^*) = 235$. Es ist also um 43 % reduziert worden.

Der konstante Faktor im Zähler der Übertragungsfunktion Gl. (2.77) beeinflusst nur die Grunddämpfung des Filters und kann daher in vielen Fällen frei gewählt werden. Wird die Nebenbedingung Gl. (2.80a) weggelassen und die Optimierung erneut durchgeführt, erhält man den optimalen Parametervektor

$$\underline{x}^* = \begin{bmatrix} 13,942 \\ 0,06268 \\ 0,06283 \\ 18,2115 \\ 0,53198 \end{bmatrix}$$

und das Empfindlichkeitsmaß $M(\underline{x}^*) = 33,7$. In diesem Fall ist es also um 92 % (oder um den Faktor 12,2) reduziert worden.

Im Bild 2.4 wird das Empfindlichkeitsverhalten des Netzwerkes vor und nach der Optimierung für beide Fälle gezeigt.

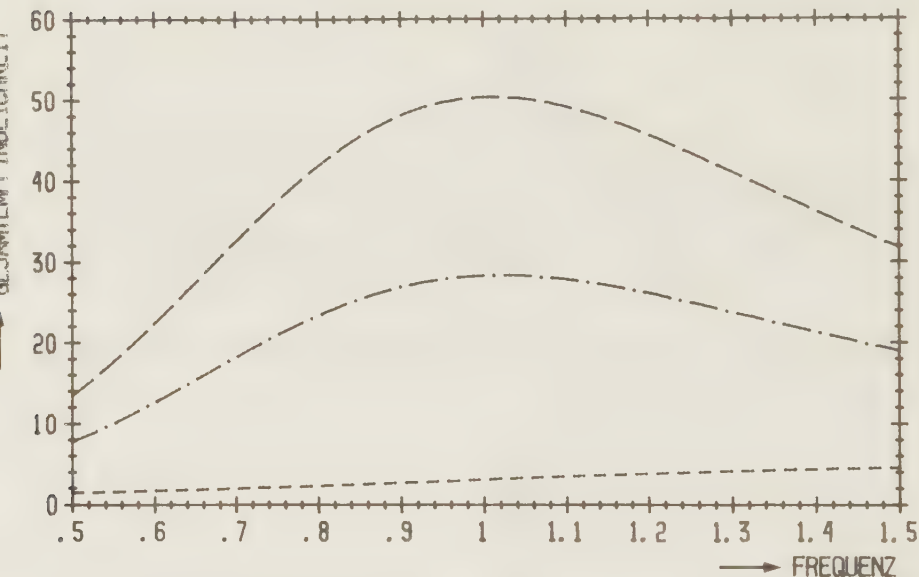


Bild 2.4: Gesamtempfindlichkeit des Netzwerkes des zweiten Beispiels

— — : vor der Optimierung

— · — : nach der Optimierung mit festem Vorfaktor

— — — : nach der Optimierung mit freiem Vorfaktor

III. TOLERANZANALYSE

1. Einführung

Beim Entwurf eines elektrischen Netzwerkes werden die Bauelementewerte im allgemeinen so festgelegt, daß die vorgeschriebene Netzwerkfunktion genau eingehalten wird oder ein gegebenes Toleranzschema bestmöglich erfüllt wird. Die dabei ermittelten Werte stellen die Sollwerte der Bauelemente dar.

Die Bauelementewerte können jedoch auf Grund fertigungsbedingter Ungenauigkeiten, Alterung oder äußerer Einflüsse (z.B. Temperatur, Feuchtigkeit, Druck) innerhalb bestimmter Toleranzgrenzen von ihren Sollwerten abweichen. Diese Schwankungen der Bauelementewerte haben Abweichungen der Netzwerkfunktion von ihrem Sollwert zur Folge.

Die Toleranzanalyse beschäftigt sich mit der Berechnung dieser Abweichungen der Netzwerkfunktion, wenn die Bauelemente innerhalb ihres Toleranzbereiches einen beliebigen Wert annehmen dürfen.

Die in der Toleranzanalyse verwendeten Verfahren können in zwei grundlegend verschiedene Gruppen aufgeteilt werden:

a) Worst-Case-Verfahren

In diesem Verfahren wird die Abweichung der Netzwerkfunktion für den ungünstigsten Fall (worst-case) aller möglichen Kombinationen der Bauelemente ermittelt.

Da die Wahrscheinlichkeit für das Auftreten der Bauelementekombination, die den ungünstigsten Fall ergeben, äußerst gering ist, wird im allgemeinen ein Auslegen der Bauelementetoleranzen für den Worst-Case unwirtschaftlich sein. Die Dimensionierung einer Schaltung nach diesem Kriterium ist daher nur dann sinnvoll, wenn absolute Zuverlässigkeit gefordert wird.

b) Statistische Verfahren

In den statistischen Verfahren werden die Bauelemente als Zufallsgrößen mit bekannter Häufigkeitsverteilung innerhalb ihres Toleranzbereichs aufgefaßt. Die Netzwerkfunktion selbst wird dadurch ebenfalls zu einer Zufallsgröße, deren Häufigkeitsverteilung berechnet werden kann. Aus dieser Verteilung lassen sich dann die Werte der Netzwerkfunktion berechnen, innerhalb derer (im statistischen Mittel) ein gewisser Anteil aller realisierten Netzwerke liegen wird. Diese Werte werden statistische Toleranzgrenzen genannt.

Im Abschnitt 2 dieses Kapitels werden die Worst-Case-Verfahren und in Abschnitt 3 die statistischen Verfahren kurz behandelt. Im Abschnitt 4 wird dann ein spezielles statistisches Verfahren, welches wesentlich genauere Ergebnisse als die bis dahin behandelten Verfahren liefert, eingehend dargelegt.

2. Worst-Case-Verfahren

2.1 Linearisierung der Netzwerkfunktion

Die reelle Netzwerkfunktion $f(\underline{x})$ wird in der Umgebung der Sollwerte der Bauelemente, $\underline{x}_0 = (x_{01}, x_{02}, \dots, x_{0n})^t$, durch eine Taylor-Reihe erster Ordnung approximiert

$$f(\underline{x}) = f(\underline{x}_0) + \sum_{i=1}^n \left. \frac{\partial f(\underline{x})}{\partial x_i} \right|_{\underline{x}_0} (x_i - x_{0i}) \quad (3.1)$$

mit $\Delta x_i = x_i - x_{0i}$ und der halbrelativen Empfindlichkeit an der Stelle $\underline{x}_0$, $S_{x_i}^f = \left. \frac{\partial f(\underline{x})}{\partial x_i / x_i} \right|_{\underline{x}_0}$ ist

$$f(\underline{x}) = f(\underline{x}_0) + \sum_{i=1}^n S_{x_i}^f \frac{\Delta x_i}{x_{0i}}. \quad (3.2)$$

Die Worst-Case-Grenzen können nun näherungsweise aus dem Maximum bzw. Minimum der linearisierten Netzwerkfunktion Gl. (3.2) berechnet werden. Die Bauelemente dürfen dabei innerhalb ihres Toleranzbereiches

$$(1-t_i)x_{0i} \leq x_i \leq (1+t_i)x_{0i} \quad i=1,2,\dots,n \quad (3.3)$$

einen beliebigen Wert einnehmen.

Aus Gl. (3.2) ist damit sofort zu erkennen, daß das Maximum genau dann auftritt, wenn alle Parameteränderungen Δx_i ($i=1,2,\dots,n$) jeweils an der Grenze ihres Toleranzbereichs liegen und ihr Vorzeichen mit dem der entsprechenden Empfindlichkeit übereinstimmt. Analog dazu tritt das Minimum dann auf, wenn Δx_i und die Empfindlichkeit $S_{x_i}^f$ entgegengesetztes Vorzeichen haben.

Die Abweichung der Netzwerkfunktion ist daher im Worst-Case

$$|\Delta f_{\max}| = \sum_{i=1}^n |S_{x_i}^f| t_i. \quad (3.4)$$

Die Nachteile dieses Verfahren für die Bestimmung des Worst-Case sind folgende [31]:

- a) Diese Methode kann nur dann angewandt werden, wenn die Linearisierung in Gl. (3.1) eine hinreichend gute Näherung liefert. Sie scheitert daher in Fällen, in denen die Terme höherer Ordnung in der Taylorentwicklung auch für kleine Toleranzen nicht mehr vernachlässigt werden können. Dies ist z.B. für die Dämpfung in den Anpassungsstellen im Durchlaßbereich von Reaktanzfiltern der Fall [32]. Hier verschwinden nämlich sämtliche Empfindlichkeiten erster Ordnung der Dämpfung exakt.
- b) Diese Methode scheitert außerdem dann, wenn Vorzeichenumkehrungen der Empfindlichkeiten innerhalb des Toleranzbereichs der Bauelemente auftreten, oder wenn eine kleine Änderung eines Elements eine zu große Änderung in der Empfindlichkeit eines anderen hervorruft.
- c) In diesem Verfahren wird davon ausgegangen, daß der ungünstigste Fall an den Grenzen der Toleranzbereiche der Bauelemente auftritt. Im allgemeinen wird jedoch das Maximum bzw. Minimum der Netzwerkfunktion an einem beliebigen Punkt im (durch die Toleranzen beschränkten) Parameterraum liegen.

2.2 Die Berechnung des Worst-Case mit Hilfe nichtlinearer Programmierungsverfahren

Die Berechnung des maximalen bzw. minimalen Wertes, den die Netzwerkfunktion im Toleranzbereich der Bauelemente annehmen kann, läßt sich auch als nichtlineares Programmierungsproblem auffassen [24], [25], [31]. Dieses kann folgendermaßen formuliert werden:

Es ist ein Parametervektor $\underline{x}^*$ so zu bestimmen, daß die Netzwerkfunktion entweder zu einem Maximum oder einem Minimum wird. Dabei sind die Nebenbedingungen

$$(1-t_i)x_{oi} \leq x_i \leq x_{oi}(1+t_i) \quad i=1,2,\dots,n$$

einzuhalten, wobei x_{oi} und t_i ($i=1,2,\dots,n$) die Sollwerte bzw. die Toleranzen der Bauelemente sind.

Dieses Optimierungsproblem mit Nebenbedingungen läßt sich z.B. mit der in Kap. II, Abschnitt 3 beschriebenen Methode von Fiacco und McCormick lösen.

Mit diesem Verfahren wird das Maximum bzw. Minimum im gesamten (beschränkten) Parameterraum und nicht nur auf dem Rand davon gesucht.

Der Vorteil dieser Methode gegenüber der im vorangegangenen Abschnitt beschriebenen liegt darin, daß man mit ihr den genauen Worst-Case, und nicht nur eine Näherung davon, ermitteln kann. Nachteilig ist hingegen der durch die Optimierung bedingte, sehr viel größere Rechenaufwand. Aus diesem Grund wird ein solches Verfahren nur in ganz speziellen Fällen sinnvoll angewandt werden können.

3. Statistische Toleranzanalyse^{*)}

In den Worst-Case-Verfahren werden die Werte der Netzwerkfunktion gesucht, die unter den ungünstigsten Bedingungen auftreten können. Die Wahrscheinlichkeit, daß diese ungünstigste Kombination der Parameterwerte auftritt, ist jedoch äußerst gering.

Anstatt das Netzwerk so zu dimensionieren, daß auch im Worst-Case die zulässigen Grenzen der Netzwerkfunktion nicht überschritten werden, was zu sehr engen Toleranzforderungen führen würde, wird es im allgemeinen wirtschaftlicher sein, wenn ein gewisser Anteil der gefertigten Schaltungen diese Grenzen überschreiten darf. Wird ein dadurch entstehender Ausschuß in Kauf genommen, kommt man nämlich zu erheblich "weicheeren" Toleranzforderungen an die Bauelemente.

In der statistischen Toleranzanalyse sind nun nicht mehr die Grenzen der Netzwerkfunktion zu berechnen, die von allen realisierten Schaltungen eingehalten werden, sondern diejenigen innerhalb derer (im statistischen Mittel) nur noch ein vorgegebener Anteil aller Schaltungen liegen wird. Hierzu wird folgendermaßen vorgegangen:

- a) Mit Hilfe eines geeigneten Verfahrens wird aus der Wahrscheinlichkeitsdichte der Bauelemente die Wahrscheinlichkeitsdichte der Netzwerkfunktion berechnet.
- b) Aus der Wahrscheinlichkeitsdichte der Netzwerkfunktion wird eine untere und eine obere Grenze berechnet, innerhalb derer der Wert der Netzwerkfunktion mit einer vorgegebenen Wahrscheinlichkeit liegen wird. Diese Grenzen werden fortan als statistische Toleranzgrenzen oder einfach Toleranzgrenzen der Netzwerkfunktion bezeichnet.

^{*)} Im Anhang wird eine Zusammenfassung der in dieser Arbeit verwendeten Begriffe der Wahrscheinlichkeitsrechnung gegeben.

Im Bild 3.1 wird ein möglicher Verlauf der Wahrscheinlichkeitsdichte der Netzwerkfunktion und die beiden Toleranzgrenzen f_u und f_o gezeigt. Die schraffierten Flächen außerhalb der Toleranzgrenzen entsprechen der sog. Ausschußrate, d.h. dem Anteil aller realisierten Schaltungen, die nicht die Toleranzforderungen erfüllen.

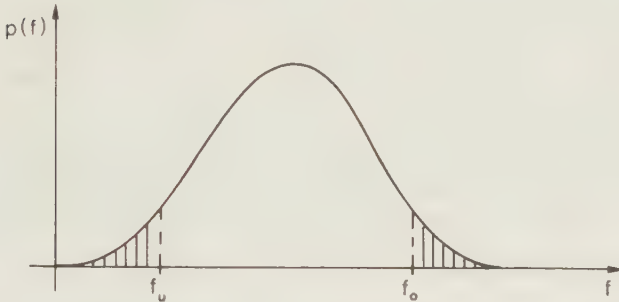


Bild 3.1: Wahrscheinlichkeitsdichtefunktion und Toleranzgrenzen

Es gibt verschiedene Methoden, mit denen die Wahrscheinlichkeitsdichte einer Netzwerkfunktion berechnet werden kann, wenn die Wahrscheinlichkeitsdichten der Bauelemente bekannt sind (siehe hierzu z.B. die Aufsätze [33|, [34|, [35|, [36|). In den folgenden beiden Abschnitten sollen davon nur zwei Verfahren, die wegen ihrer praktischen Anwendbarkeit von Interesse sind, kurz behandelt werden. Es sind diese das Verfahren nach dem Fehlerfortpflanzungsgesetz und das Monte-Carlo-Verfahren. Im Abschnitt 4 dieses Kapitels wird dann ein weiteres Verfahren eingehend behandelt, das genauer als das Verfahren nach dem Fehlerfortpflanzungsgesetz und schneller als das Monte-Carlo-Verfahren ist.

3.1 Das Verfahren nach dem Fehlerfortpflanzungsgesetz

Dieses ist das Verfahren, welches am häufigsten in der statistischen Toleranzanalyse von Netzwerken angewendet wird. Es be-

ruht auf dem zentralen Grenzwertsatz der Statistik, der besagt, daß die Verteilung einer Summe von n statistisch voneinander unabhängigen Zufallsgrößen mit endlichem Mittelwert und Streuung für $n \rightarrow \infty$ zu einer Normalverteilung tendiert [37]. Dieser Satz kann daher angewendet werden, wenn die Netzwerkfunktion hinreichend genau durch eine Taylorreihe erster Ordnung beschrieben werden kann, d.h. wenn

$$\Delta f = \sum_{i=1}^n \frac{\partial f(\underline{x})}{\partial x_i} \Delta x_i \quad (3.5)$$

ist, wenn die Anzahl der Bauelemente groß ist und wenn die Ableitungen der Netzwerkfunktion annähernd die gleiche Größenordnung haben. Die Form der Bauelementeverteilungen spielt dabei keine Rolle. Wichtig sind lediglich deren Mittelwert und Streuung.

Wird nun unter den o.g. Voraussetzungen die Netzwerkfunktion als normalverteilt angenommen, genügt für die vollständige Beschreibung der Verteilung die Berechnung von Mittelwert und Streuung (oder Varianz) der Netzwerkfunktion.

Mit Gl. (A.28) ergibt sich für den Mittelwert der Änderung der Netzwerkfunktion nach Gl. (3.5)

$$\overline{\Delta f} = E(\Delta f) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} E(\Delta x_i) \quad (3.6)$$

Mit Hilfe der Gl. (A.12) läßt sich die Varianz aus der Beziehung

$$\sigma_{\Delta f}^2 = E(\Delta f^2) - (E(\Delta f))^2 \quad (3.7)$$

berechnen. Durch Quadrieren von Gl. (3.5) und Bilden des Erwartungssatzes ergibt sich

$$E(\Delta f^2) = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right)^2 E(\Delta x_i^2) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} E(\Delta x_i \Delta x_j) \quad (3.8)$$

Durch Quadrieren von Gl. (3.6) erhält man andererseits

$$\{E(\Delta f)\}^2 = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}\right)^2 \{E(\Delta x_i)\}^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} E(\Delta x_i) E(\Delta x_j) . \quad (3.9)$$

Durch Einsetzen der Gln. (3.8) und (3.9) in Gl. (3.7) folgt

$$\sigma_{\Delta f}^2 = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}\right)^2 \{E(\Delta x_i^2) - \{E(\Delta x_i)\}^2\} + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \{E(\Delta x_i \Delta x_j) - E(\Delta x_i) E(\Delta x_j)\}$$

und auf Grund der Gleichungen (A.12) und (A.33) für die Varianz und Kovarianz ist

$$\sigma_{\Delta f}^2 = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_{x_i}^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \text{cov}(x_i, x_j) . \quad (3.10)$$

Da voraussetzungsgemäß die Bauelemente statistisch unabhängig voneinander sind, sind die Kovarianzen alle gleich Null, womit die Doppelsumme auf der rechten Seite von Gl. (3.10) verschwindet. Es gilt also

$$\sigma_{\Delta f}^2 = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_{x_i}^2 . \quad (3.11)$$

Diese Gleichung wird nach Gauß als Fehlerfortpflanzungsgesetz bezeichnet.

Mit den nach Gln. (3.6) und (3.11) berechneten Werten von Mittelwert und Varianz ist die Normalverteilung eindeutig definiert:

$$p(\Delta f) = \frac{1}{\sqrt{2\pi}\sigma_{\Delta f}} e^{-\frac{(\Delta f - \overline{\Delta f})^2}{2\sigma_{\Delta f}^2}} . \quad (3.12)$$

In den meisten Fällen sind die Verteilungen der Bauelemente symmetrisch. Dann ist $E(\Delta x_i) = 0$ für $i = 1, 2, \dots, n$ und wegen Gl. (3.6) $\overline{\Delta f} = 0$.

Mit Hilfe spezieller Kurven oder Tabellen, wie sie z.B. in [44]

gegeben sind, können für die Normalverteilung die Toleranzgrenzen der Netzwerkfunktion ermittelt werden. Häufig werden als Toleranzgrenzen die Werte $\bar{f} \pm \sigma_f$, $\bar{f} \pm 2\sigma_f$ oder $\bar{f} \pm 3\sigma_f$ verwendet. Innerhalb dieser Grenzen liegen dann 68,3 % bzw. 95,5 % oder 99,7 % aller realisierten Schaltungen.

Das Fehlerfortpflanzungsgesetz kann nur dann sinnvoll angewandt werden, wenn die am Anfang des Abschnittes erwähnten Bedingungen erfüllt sind. Sobald die Ableitungen zweiter und höherer Ordnung gegenüber denen erster Ordnung in der Taylorentwicklung nicht mehr vernachlässigt werden können und die einzelnen Terme der Summe in Gl. (3.5) nicht die gleiche Größenordnung haben, kann nicht mehr erwartet werden, daß die Netzwerkfunktion $f(\underline{x})$ normalverteilt ist. In diesem Fall genügen Mittelwert und Varianz allein nicht für die vollständige Beschreibung der Verteilung.

3.2 Das Monte-Carlo-Verfahren

Das Monte-Carlo-Verfahren beruht auf der Simulation einer großen Zahl von Schaltungen, deren Bauelementewerte durch einen Zufallszahlengenerator gemäß einer vorgeschriebenen Verteilung erzeugt werden. Die Berechnung der Netzwerkfunktion wird einige 100- oder 1000-mal mit den zufällig ausgewählten Bauelementewerten wiederholt. Aus den in Klassen eingeordneten Werten läßt sich ein Histogramm konstruieren, welches eine Schätzung der gesuchten Verteilung der Netzwerkfunktion darstellt.

"Selbst eine große Zahl von Versuchen (z.B. 10 000) kann nicht darüber hinwegtäuschen, daß es sich hierbei nur um Stichproben aus einer unendlichen Gesamtheit handelt. Als solche sind sie deshalb mit einer statistischen Unsicherheit versehen, die von der Zahl der Versuche abhängt." |33|

In der statistischen Toleranzanalyse interessieren an erster Stelle die Toleranzgrenzen der Netzwerkfunktion und nicht so sehr die Verteilung selbst. Da in der Monte-Carlo-Analyse die

Netzwerkfunktion nicht, wie im Verfahren nach dem Fehlerfortpflanzungsgesetz, approximiert wird, treten auch die dadurch bedingten und im Abschnitt 3.1 erwähnten Nachteile nicht auf. Die Ermittlung der Toleranzgrenzen auf Grund des Histogramms liefert jedoch sehr ungenaue Werte, da bei einer zu geringen Anzahl von Analysen die Ränder der Verteilung sehr schlecht bestimmt sind. Dieser Nachteil könnte mit einer Erhöhung der Anzahl von Analysen behoben werden. Um den damit verbundenen hohen Aufwand an Rechenzeit zu vermeiden, wird in [39] vorgeschlagen, das Histogramm durch eine stetige Verteilung zu approximieren, aus der dann die Toleranzgrenzen berechnet werden können.

4. Die Näherung der Wahrscheinlichkeitsdichte durch Pearsonsche Kurven

Wie bereits im Abschnitt 3.1 erwähnt, läßt sich eine Netzwerkfunktion nicht immer hinreichend genau durch eine Taylorreihe erster Ordnung approximieren. In solchen Fällen müssen noch die Glieder zweiter und evtl. höherer Ordnung hinzugezogen werden. Dies bedeutet jedoch, daß der zentrale Grenzwertsatz der Statistik nicht mehr angewendet werden kann, da nun die Netzwerkfunktion außer aus Summen der Zufallsvariablen auch aus Produkten dieser besteht. Es kann daher nicht mehr erwartet werden, daß die Netzwerkfunktion normalverteilt ist und sich somit allein durch ihren Mittelwert und Streuung beschreiben läßt.*)

Für die genauere Beschreibung der Verteilung wird daher die Berechnung von Momenten höherer Ordnung notwendig sein. In den meisten Fällen ist eine Charakterisierung durch die Momente bis zur vierten Ordnung ausreichend [39]. Für die Approximation der Verteilung der Netzwerkfunktion eignet sich besonders gut das Pearsonsche Kurvensystem [39], [40], [41], da in diesem System die Kurven von vier Parametern abhängen, die aus den Momenten bis zur vierten Ordnung berechnet werden können.

Das ursprünglich von Pearson vorgesehene Aufgabengebiet für dieses Kurvensystem war die Näherung einer empirischen Häufigkeitsverteilung durch eine stetige Verteilung. Dies entspricht z.B. der im Abschnitt 3.2 erwähnten Näherung des Histogramms aus der Monte-Carlo-Analyse durch eine stetige Verteilung. In einem solchen Fall werden aus dem Histogramm erst die Momente bis zur vierten Ordnung berechnet, aus denen dann die Parameter der Pearsonschen Kurve gewonnen werden können.

*) "Bei der Bildung des Produkts von unabhängigen Zufallsgrößen geht die Produktgröße asymptotisch gegen eine 'Logarithmus-Normalverteilung' und nicht, wie aus dem Verfahren der Fehlerfortpflanzung geschlossen werden könnte, gegen eine Normalverteilung." [35]

Im Gegensatz hierzu werden im folgenden für die Toleranzanalyse die Momente der Netzwerkfunktion aus der Taylor-Entwicklung zweiter Ordnung berechnet. Aus den Momenten bis zur vierten Ordnung werden dann erst Schiefe und Exzeß der Verteilung ermittelt. Je nach den Werten dieser beiden Parameter wird ein spezieller Kurventyp aus dem Pearsonschen Kurvensystem ausgewählt. Einer dieser Kurventypen (Typ I) wurde bereits in [39] für die Näherung der Wahrscheinlichkeitsdichte von Netzwerkfunktionen verwendet, deren Schiefe und Exzeß auf einen bestimmten Wertebereich beschränkt ist. Die Verteilung kann jedoch häufig nicht durch eine Pearson-Kurve vom Typ I dargestellt werden. Aus diesem Grund werden hier auch noch die anderen Kurventypen hinzugezogen, so daß sich Wahrscheinlichkeitsdichten mit beliebiger Schiefe und Exzeß darstellen lassen.

4.1 Berechnung der Momente der Netzwerkfunktion

Werden in der Darstellung der Änderung der Netzwerkfunktion gemäß Gl. (3.5) die Terme zweiter Ordnung hinzugezogen, dann gilt

$$\Delta f = \sum_{i=1}^n \frac{\partial f(\underline{x})}{\partial x_i} \Delta x_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j} \Delta x_i \Delta x_j, \quad (3.13)$$

wobei die Ableitungen der Netzwerkfunktion $f(\underline{x})$ für die Sollwerte der Bauelemente x_{oi} ($i=1,2,\dots,n$) zu berechnen sind. Hieraus sollen nun die Momente bis zur vierten Ordnung berechnet werden. Es soll dabei vorausgesetzt werden, daß die Bauelemente statistisch unabhängig sind, eine symmetrische Wahrscheinlichkeitsdichte besitzen und ihr Sollwert mit dem Mittelwert übereinstimmt. Die Symmetrie bewirkt, daß alle Momente ungerader Ordnung von Δx_i verschwinden, d.h.

$$E(\Delta x_i^m) = 0 \quad \text{für } i = 1, 2, \dots, n \quad (3.14) \\ \text{und } m = 1, 3, 5, 7, \dots$$

Wegen der statistischen Unabhängigkeit ist, gemäß Gl. (A.26),

$$E(\Delta x_i \Delta x_j) = E(\Delta x_i) E(\Delta x_j) \quad \text{für alle } i \neq j$$

und wegen Gl. (3.14)

$$E(\Delta x_i \Delta x_j) = 0 \quad \text{für alle } i \neq j. \quad (3.15)$$

Für $i = j$ ist jedoch, da der Mittelwert mit dem Sollwert übereinstimmt,

$$E(\Delta x_i^2) = E(x_i - E(x_i))^2 \quad (3.16)$$

und daher

$$\sigma_{\Delta x_i}^2 = \sigma_{x_i}^2. \quad (3.17)$$

Mittelwert oder Momente erster Ordnung

Durch Anwendung der Gl. (A.28) auf Gl. (3.13) ergibt sich für den Mittelwert von Δf

$$m_1 = E(\Delta f) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} E(\Delta x_i) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} E(\Delta x_i \Delta x_j)$$

und auf Grund der Gln. (3.14) und (3.15)

$$m_1 = E(\Delta f) = \frac{1}{2} \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2} E(\Delta x_i^2). \quad (3.18)$$

Moment zweiter Ordnung

Durch Quadrieren von Gl. (3.13) erhält man

$$\begin{aligned} \Delta f^2 &= \sum_{i=1}^n \sum_{j=1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \Delta x_i \Delta x_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \frac{\partial f}{\partial x_i} \frac{\partial^2 f}{\partial x_j \partial x_k} \Delta x_i \Delta x_j \Delta x_k \\ &+ \frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} \frac{\partial^2 f}{\partial x_k \partial x_l} \Delta x_i \Delta x_j \Delta x_k \Delta x_l. \end{aligned}$$

Unter Berücksichtigung der Gln. (A.29) und (3.14) verschwinden bei der Bildung des Erwartungswertes alle Terme, in denen ein

Faktor Δx_i^m mit ungeradem m enthalten ist. Daher ist

$$\begin{aligned} m_2 = E(\Delta f^2) &= \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right)^2 E(\Delta x_i^2) + \frac{1}{4} \sum_{i=1}^n \left(\frac{\partial^2 f}{\partial x_i^2} \right)^2 E(\Delta x_i^4) \\ &+ \frac{1}{4} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\partial^2 f}{\partial x_i^2} \frac{\partial^2 f}{\partial x_j^2} E(\Delta x_i^2) E(\Delta x_j^2) \\ &+ \frac{1}{2} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)^2 E(\Delta x_i^2) E(\Delta x_j^2) \end{aligned}$$

oder

$$\begin{aligned} m_2 &= \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right)^2 E(\Delta x_i^2) + \frac{1}{4} \sum_{i=1}^n \left(\frac{\partial^2 f}{\partial x_i^2} \right)^2 E(\Delta x_i^4) \\ &+ \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left\{ \frac{1}{2} \frac{\partial^2 f}{\partial x_i^2} \frac{\partial^2 f}{\partial x_j^2} + \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)^2 \right\} E(\Delta x_i^2) E(\Delta x_j^2) \end{aligned} \quad (3.19)$$

Zwecks Vereinfachung der Gln. (3.18) und (3.19) und der noch zu berechnenden Formeln für die Momente dritter und vierter Ordnung, werden folgende Beziehungen eingeführt

$$a_i = x_{oi} \frac{\partial f}{\partial x_i} = S_{x_i}^f \quad (i = 1, 2, \dots, n) \quad (3.20a)$$

$$b_{ii} = \frac{1}{2} x_{oi}^2 \frac{\partial^2 f}{\partial x_i^2} = \frac{1}{2} S_{x_i x_i}^f \quad (i=1, 2, \dots, n) \quad (3.20b)$$

$$b_{ij} = x_{oi} x_{oj} \frac{\partial^2 f}{\partial x_i \partial x_j} = S_{x_i x_j}^f \quad \begin{matrix} (i=1, 2, \dots, n-1 \\ j=i+1, \dots, n) \end{matrix} \quad (3.20c)$$

Es sei außerdem

$$e_i = \frac{\Delta x_i}{x_{oi}} \quad (i = 1, 2, \dots, n) \quad (3.21)$$

die relative Abweichung der Bauelemente von ihrem Sollwert. Werden die Gln. (3.20a-c) und (3.21) in die Gln. (3.18) und (3.19) eingeführt, erhält man

$$m_1 = E(\Delta f) = \sum_{i=1}^n b_{ii} E(e_i^2) \quad (3.22)$$

$$m_2 = E(\Delta f^2) = \sum_{i=1}^n a_i^2 E(e_i^2) + \sum_{i=1}^n b_{ii}^2 E(e_i^4) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n (2 b_{ii} b_{jj} + b_{ij}^2) E(e_i^2) E(e_j^2). \quad (3.23)$$

Momente dritter und vierter Ordnung

Ähnlich wie das Moment zweiter Ordnung können nun auch die Momente dritter und vierter Ordnung berechnet werden. Auf Grund des relativ hohen Aufwandes an algebraischer Manipulation sollen hier nur die Ergebnisse wiedergegeben werden [42]:

$$\begin{aligned} m_3 = E(\Delta f^3) &= \sum_{i=1}^n C_i^{(4)} E(e_i^4) + \sum_{i=1}^n C_i^{(6)} E(e_i^6) + \\ &+ \sum_{i=1}^{n-1} \sum_{j=i+1}^n 4 C_{ij}^{(22)} E(e_i^2) E(e_j^2) + \\ &+ C_{ij}^{(42)} E(e_i^4) E(e_j^2) + C_{ij}^{(24)} E(e_i^2) E(e_j^4) + \\ &+ \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n C_{ijk}^{(222)} E(e_i^2) E(e_j^2) E(e_k^2) \end{aligned} \quad (3.24)$$

$$C_i^{(4)} = 3 a_i^2 b_{ii}, C_i^{(6)} = b_{ii}^3, C_{ij}^{(22)} = 3(a_i^2 b_{ij} +$$

$$+ a_j^2 b_{ii} + 2 a_i a_j b_{ij})$$

$$C_{ij}^{(42)} = 3(b_{ii} b_{ij}^2 + b_{jj} b_{ii}^2), C_{ij}^{(24)} = 3(b_{jj} b_{ij}^2 + b_{ii} b_{jj}^2)$$

$$C_{ijk}^{(222)} = 3(b_{ii} b_{jk}^2 + b_{kk} b_{ij}^2 + b_{jj} b_{ik}^2 +$$

$$+ 2(b_{ii} b_{jj} b_{kk} + b_{ij} b_{jk} b_{ik}))$$

$$\begin{aligned}
 m_4 = E(\Delta f^4) = & \sum_{i=1}^n D_i^{(4)} E(e_i^4) + D_i^{(6)} E(e_i^6) + D_i^{(8)} E(e_i^8) + \\
 & + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \{ D_{ij}^{(22)} E(e_i^2) E(e_j^2) + \\
 & + D_{ij}^{(42)} E(e_i^4) E(e_j^2) + D_{ij}^{(24)} E(e_i^2) E(e_j^4) + \\
 & + D_{ij}^{(62)} E(e_i^6) E(e_j^2) + D_{ij}^{(26)} E(e_i^2) E(e_j^6) + \\
 & + D_{ij}^{(44)} E(e_i^4) E(e_j^4) \} + \\
 & + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n \{ D_{ijk}^{(222)} E(e_i^2) E(e_j^2) E(e_k^2) + \\
 & + D_{ijk}^{(422)} E(e_i^4) E(e_j^2) E(e_k^2) + \\
 & + D_{ijk}^{(242)} E(e_i^2) E(e_j^4) E(e_k^2) + \\
 & + D_{ijk}^{(224)} E(e_i^2) E(e_j^2) E(e_k^4) \} + \\
 & + \sum_{i=1}^{n-3} \sum_{j=i+1}^{n-2} \sum_{k=j+1}^{n-1} \sum_{l=k+1}^n D_{ijkl}^{(2222)} E(e_j^2) E(e_j^2) E(e_k^2) \\
 & E(e_l^2) \\
 & (3.25)
 \end{aligned}$$

$$D_i^{(4)} = a_i^4, D_i^{(6)} = 6a_i^2 b_{ii}^2, D_i^{(8)} = b_{ii}^4, D_{ij}^{(22)} = 6a_i^2 a_j^2$$

$$D_{ij}^{(42)} = 6a_i^2 b_{ij}^2 + 6b_{ii}^2 a_j^2 + 12a_i^2 b_{ii} b_{jj} + 24a_i a_j b_{ii} b_{ij}$$

$$D_{ij}^{(24)} = 6a_j^2 b_{ij}^2 + 6b_{jj}^2 a_i^2 + 12a_j^2 b_{ii} b_{jj} + 24a_i a_j b_{jj} b_{ij}$$

$$\begin{aligned}
 D_{ijk}^{(222)} = & 6(a_i^2 b_{jk}^2 + a_j^2 b_{ik}^2 + a_k^2 b_{ij}^2) + \\
 & + 12(a_i^2 b_{jj} b_{kk} + a_j^2 b_{kk} b_{ii} + a_k^2 b_{ii} b_{jj}) +
 \end{aligned}$$

$$+ 24(a_i a_j (b_{ik} b_{jk} + b_{ij} b_{kk}) + a_j a_k (b_{ij} b_{ik} + b_{jk} b_{ii}) + a_i a_k (b_{ij} b_{jk} + b_{jj} b_{ik}))$$

$$D_{ij}^{(62)} = 2b_{ii}^2 (3b_{ij}^2 + 2b_{ii} b_{jj}),$$

$$D_{ij}^{(26)} = 2b_{jj}^2 (3b_{ij}^2 + 2b_{jj} b_{ii})$$

$$D_{ij}^{(44)} = b_{ij}^4 + 6b_{ii}^2 b_{jj}^2 + 12b_{ii} b_{jj} b_{ij}^2$$

$$D_{ijk}^{(422)} = 6b_{ii}^2 b_{jk}^2 + 6b_{ij}^2 b_{ik}^2 + 12b_{ii} b_{jj} b_{ik}^2 + \\ + 12b_{ii} b_{kk} b_{ij}^2 + 12b_{jj} b_{kk} b_{ii}^2 + 24b_{ii} b_{ij} b_{ik} b_{jk}$$

$$D_{ijk}^{(242)} = 6b_{jj}^2 b_{ik}^2 + 6b_{ij}^2 b_{jk}^2 + 12b_{jj} b_{kk} b_{ij}^2 + \\ + 12b_{jj} b_{ii} b_{jk}^2 + 12b_{kk} b_{ii} b_{jj}^2 + \\ + 24b_{jj} b_{jk} b_{ij} b_{ik}$$

$$D_{ijk}^{(224)} = 6b_{kk}^2 b_{ij}^2 + 6b_{jk}^2 b_{ik}^2 + 12b_{kk} b_{ii} b_{jk}^2 + \\ + 12b_{kk} b_{jj} b_{ik}^2 + 12b_{ii} b_{jj} b_{kk}^2 + \\ + 24b_{kk} b_{ik} b_{jk} b_{ij}$$

$$D_{ijkl}^{(2222)} = 6(b_{ij}^2 b_{kl}^2 + b_{ik}^2 b_{jl}^2 + b_{il}^2 b_{jk}^2) + \\ + 12(b_{ij}^2 b_{kk}^2 b_{ll} + b_{ik}^2 b_{jj} b_{ll} + b_{il}^2 b_{jj} b_{kk} + \\ + b_{jk}^2 b_{ii} b_{ll} + b_{jl}^2 b_{ii} b_{kk} + b_{kl}^2 b_{ii} b_{jj}) + \\ + 24(b_{ii} b_{jk} b_{jl} b_{kl} + b_{jj} b_{ik} b_{il} b_{kl} +$$

$$+ b_{kk}b_{ij}b_{il}b_{jl} + b_{ll}b_{ij}b_{ik}b_{jk} +$$

$$+ b_{ij}b_{jk}b_{il}b_{kl} + b_{ij}b_{ik}b_{jl}b_{kl} + b_{jk}b_{ik}b_{il}b_{jl} \cdot$$

4.2 Das Pearsonsche Kurvensystem |40|, |41|

Die Kurven des Pearsonschen Kurvensystems können aus den Lösungen der Differentialgleichung

$$\frac{dp(z)}{dz} = p(z) \frac{z+a}{B(z)} \quad (3.26)$$

generiert werden. Hierin ist $p(z)$ die Wahrscheinlichkeitsdichtefunktion der Zufallsvariablen Z und $B(z)$ ein Polynom, das der Kurvenformung dient. Die Lösung der Differentialgleichung führt in Abhängigkeit von a und der Koeffizienten des Polynoms $B(z)$ zu einer Vielzahl von Kurventypen, unter denen auch die Normalverteilung, die Betaverteilung, die Gammaverteilung und die Logarithmus-Normalverteilung zu finden sind.

Im folgenden sollen die Koeffizienten des Polynoms $B(z) = b_0 + b_1z + b_2z^2 + \dots$ und der Parameter a in Abhängigkeit der zentralen Momente der zu approximierenden Verteilung berechnet werden. Hierzu wird der Einfachheit halber der Ursprung des Koordinatensystems mit dem Mittelwert der Verteilung zusammengelegt. Dies bewirkt, daß die zentralen Momente mit den Momenten um den Ursprung gleich sind.

Gl. (3.26) wird zunächst in die Form

$$(b_0 + b_1z + b_2z^2 + \dots) p'(z) = p(z) \cdot (z+a)$$

gebracht. Anschließend werden beide Seiten mit z^k multipliziert und nach z integriert:

$$\int_{-\infty}^{\infty} z^k (b_0 + b_1z + b_2z^2 + \dots) p'(z) dz = \int_{-\infty}^{\infty} p(z) (z+a) z^k dz \cdot$$

Durch partielle Integration der linken Seite ist

$$z^k (b_0 + b_1 z + b_2 z^2 + \dots) p(z) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} (k b_0 z^{k-1} + (k+1) b_1 z^k + \\ + (k+2) b_2 z^{k+1} + \dots) p(z) dz = \int_{-\infty}^{\infty} z^{k+1} p(z) dz + a \int_{-\infty}^{\infty} z^k p(z) dz .$$

Bei Anwendung der Definitionsgleichung der zentralen Momente, Gl. (A.10),

$$\mu_k = \int_{-\infty}^{\infty} (z - m_1)^k p(z) dz$$

(wobei $m_1 = 0$ gilt, da der Ursprung des Koordinatensystems voraussetzungsgemäß mit dem Mittelwert identisch ist) und unter der Voraussetzung, daß der Ausdruck $z^k (b_0 + b_1 z + b_2 z^2 + \dots) p(z)$ für $z \rightarrow \pm \infty$ verschwindet, ist

$$-k b_0 \mu_{k-1} - (k+1) b_1 \mu_k - (k+2) b_2 \mu_{k+1} - \dots = \mu_{k+1} + a \mu_k$$

oder

$$a \mu_k + k b_0 \mu_{k-1} + (k+1) b_1 \mu_k + (k+2) b_2 \mu_{k+1} + \dots = -\mu_{k+1} . \quad (3.27)$$

Für $k = 0, 1, 2, \dots, s$ erhält man aus Gl. (3.27) ein lineares Gleichungssystem mit $s+1$ Gleichungen, aus denen die Koeffizienten des Polynoms $B(z)$ und des Parameters a aus Gl. (3.26) in Abhängigkeit der ersten $s+1$ zentralen Momente der Verteilung berechnet werden können.

Für den Spezialfall, daß $s=3$ ist, erhält man unter Berücksichtigung, daß $\mu_1 = 0$ (vgl. Gl. (A.11a)) und $\mu_0 = 1$ gilt, folgendes Gleichungssystem für b_0, b_1, b_2 und a :

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 3\mu_2 \\ \mu_2 & 0 & 3\mu_2 & 4\mu_3 \\ \mu_3 & 3\mu_2 & 4\mu_3 & 5\mu_4 \end{bmatrix} \begin{bmatrix} a \\ b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -\mu_2 \\ -\mu_3 \\ -\mu_4 \end{bmatrix} \quad (3.28)$$

Hier wurde also ein Polynom zweiten Grades vorausgesetzt. Die Differentialgleichung (3.26) ist daher

$$\frac{dp(z)}{dz} = p(z) \frac{z+a}{b_0+b_1z+b_2z^2} . \quad (3.29)$$

Es können auch Polynome von höherem Grad als den zweiten für die Kurvenformung verwendet werden. Für die Bestimmung der zusätzlichen Koeffizienten $b_3, b_4, \dots$ muß lediglich das Gleichungssystem (3.28) mit weiteren Gleichungen für $k = 4, 5, \dots$ ergänzt werden. Dies setzt jedoch voraus, daß auch die zentralen Momente fünfter und höherer Ordnung bekannt sind. In den meisten Fällen ist aber die Charakterisierung der Verteilung mit den Momenten bis zur vierten Ordnung ausreichend [39]. Aus diesem Grund wird auch das Polynom $B(z)$ maximal vom zweiten Grad sein können. Im folgenden wird daher immer von der Differentialgleichung (3.29) und von dem zugehörigen Gleichungssystem (3.28) ausgegangen.

Aus der Lösung des Gleichungssystems (3.28) erhält man

$$a = \frac{1}{D} \mu_3 (\mu_4 + 3\mu_2^2) , \quad (3.30a)$$

$$b_0 = -\frac{1}{D} \mu_2 (4\mu_2\mu_4 - 3\mu_3^2) , \quad (3.30b)$$

$$b_1 = -a , \quad (3.30c)$$

$$b_2 = -\frac{1}{D} (2\mu_2\mu_4 - 3\mu_3^2 - 6\mu_2^3) , \quad (3.30d)$$

$$D = 10\mu_2\mu_4 - 18\mu_2^3 - 12\mu_3^2 . \quad (3.30e)$$

Die Gln. (3.30a-e) können auch in Abhängigkeit der Schiefe

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} \quad (3.31)$$

und des Exzesses

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \quad (3.32)$$

wie folgt ausgedrückt werden:

$$a = \frac{1}{D} \frac{\mu_3}{\mu_2} (\beta_2 + 3) , \quad (3.33a)$$

$$b_0 = -\frac{1}{D'} v_2 (4s_2 - 3s_1) , \quad (3.33b)$$

$$b_1 = -a , \quad (3.33c)$$

$$b_2 = -\frac{1}{D'} (2s_2 - 3s_1 - 6) , \quad (3.33d)$$

$$D' = 2(5s_2 - 6s_1 - 9) . \quad (3.33e)$$

Nachdem die Koeffizienten a , b_0 , b_1 und b_2 auf Grund der Gln. (3.30a-e) oder (3.33a-e) aus den Momenten der Zufallsvariablen Z bekannt sind, kann durch Integration der Differentialgleichung (3.29) die Wahrscheinlichkeitsdichte $p(z)$ ermittelt werden. Je nach Art der Nullstellen des Polynoms $B(z) = b_0 + b_1 z + b_2 z^2$ wird man dabei unterschiedliche Lösungen für $p(z)$ erhalten.

Die Art der Nullstellen

$$z_{1,2} = -\frac{b_1}{2b_2} \pm \frac{1}{2b_2} \sqrt{b_1^2 - 4b_0 b_2} \quad (3.34)$$

ist von

$$K = \frac{b_1^2}{4b_0 b_2} \quad (3.35)$$

abhängig. K soll daher als Kriterium für die Unterscheidung der verschiedenen Lösungen dienen.

Die Pearsonschen Kurven können nun in die folgenden drei Grundtypen aufgeteilt werden (es werden hier die ursprünglichen Typenbezeichnungen von Pearson beibehalten):

Typ I:	$K < 0$	$B(z)$ hat reelle Nullstellen mit unterschiedlichen Vorzeichen,
Typ IV:	$0 < K < 1$	$B(z)$ hat zwei komplex konjugierte Nullstellen,
Typ VI:	$K > 1$	$B(z)$ hat zwei reelle Nullstellen mit gleichen Vorzeichen.

Für jeden dieser Typen liefert die Integration der Differenzialgleichung (3.29) eine andere Lösung $p(z)$. Neben diesen Grundtypen gibt es weitere Kurven, die an den Grenzen der jeweiligen Gebiete oder im Übergang von einem Typ zur anderen auftreten. Diese Kurventypen können zum Teil als Sonderfälle der drei Grundtypen behandelt werden.

Werden nun die Gln. (3.33a-e) in die Gl. (3.35) eingesetzt, erhält man eine Darstellung des Kriteriums K als Funktion von Schiefe und Exzeß der Verteilung

$$K = \frac{\beta_1 (\beta_2 + 3)^2}{4(2\beta_2 - 3\beta_1 - 6)(4\beta_2 - 3\beta_1)} \quad (3.36)$$

Aus dieser Beziehung und dem zulässigen Wertebereich von K lassen sich Bedingungsgleichungen für β_1 und β_2 für jeden der drei Grundtypen angeben. Werden diese Bedingungsgleichungen in ein (β_1, β_2) -Diagramm eingetragen, erhält man für jeden Kurventyp ein zulässiges Gebiet. Sind Schiefe und Exzeß einer Zufallsvariable bekannt, kann aus diesem Diagramm sofort ermittelt werden, mit welchem Kurventyp des Pearsonschen Kurvensystems die Verteilung zu approximieren ist. Bild 3.2 zeigt das (β_1, β_2) -Diagramm und die Regionen der drei Grundtypen. In dem Diagramm erscheinen außerdem auch die Sondertypen des Kurvensystems.

Die Gebiete der drei Grundtypen lassen sich wie folgt abstecken:

- a) Typ I: Da für Typ I $K < 0$ gilt, muß wegen $\beta_1 \geq 0$ (vgl. Gl. (3.31)) in Gl. (3.36)

$$(2\beta_2 - 3\beta_1 - 6)(4\beta_2 - 3\beta_1) < 0 \quad (3.37a)$$

sein. Hiermit wird im (β_1, β_2) -Diagramm ein Gebiet abgesteckt, das zwischen den beiden Geraden

$$2\beta_2 - 3\beta_1 - 6 = 0 \quad (3.37b)$$

und $4\beta_2 - 3\beta_1 = 0 \quad (3.37c)$

und der Koordinatenachse $\beta_1 = 0$ liegt. Wie noch bei der ausführ-

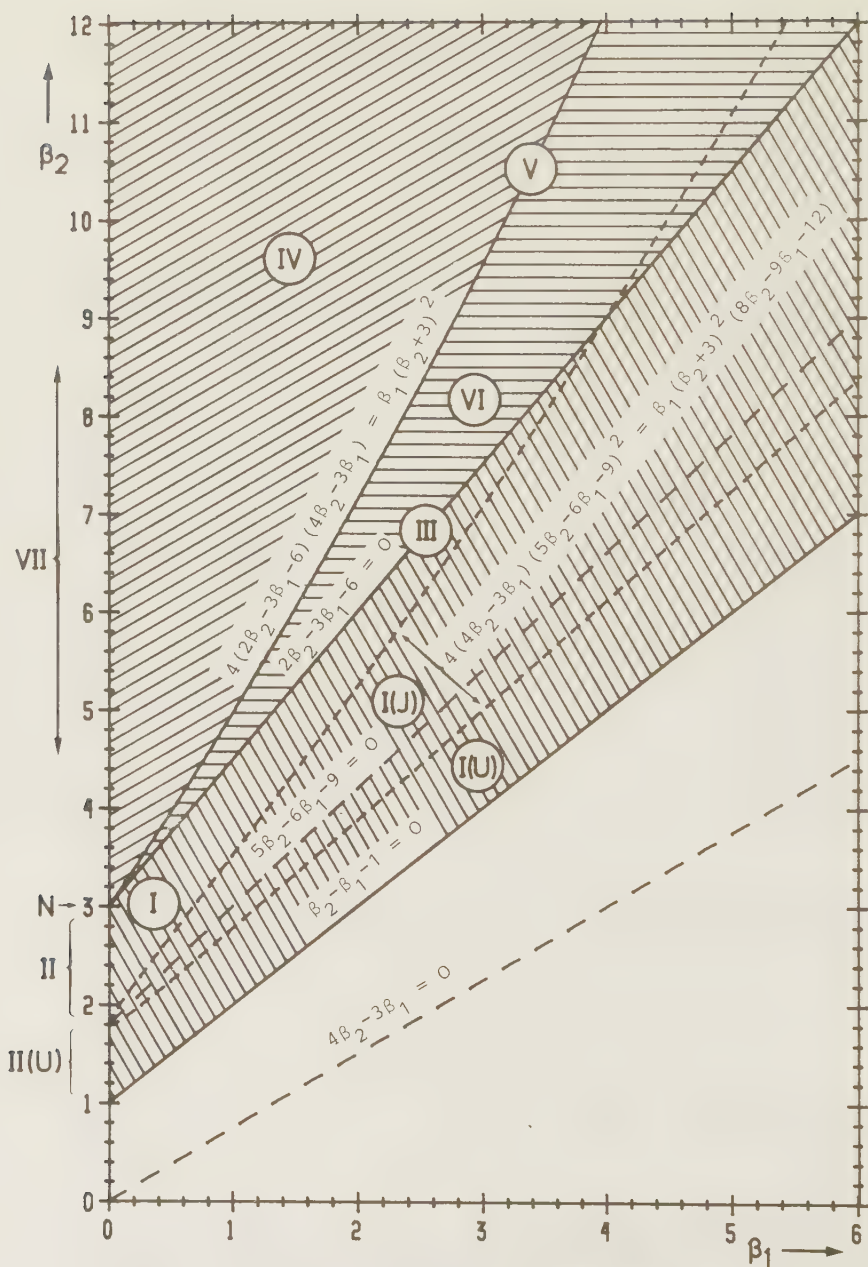


Bild 3.2: Regionen im (β_1, β_2) -Diagramm für die verschiedenen Pearsonschen Kurven

licheren Behandlung des Kurventyps I in Abschnitt 4.2.1 gezeigt werden wird, ist die absolute untere Grenze für Schiefe und Exzeß jedoch bereits durch die Gerade $\beta_2 - \beta_1 - 1 = 0$ gegeben.

b) Typ IV: Da für diesen Typ $0 < K < 1$ gilt, wird, wegen $\beta_1 \geq 0$, der zulässige Bereich durch die Ungleichung

$$\infty > 4(2\beta_2 - 3\beta_1 - 6)(4\beta_2 - 3\beta_1) > \beta_1(\beta_2 + 3)^2 \quad (3.38a)$$

definiert. Das Gebiet der Pearsonschen Kurven vom Typ IV wird daher, wie im Bild 3.2 dargestellt, außer durch die Achse $\beta_1 = 0$ nur noch durch die Kurve

$$4(2\beta_2 - 3\beta_1 - 6)(4\beta_2 - 3\beta_1) = \beta_1(\beta_2 + 3)^2 \quad (3.38b)$$

begrenzt.

c) Typ VI: Damit in Gl. (3.36) die Bedingung $K > 1$ erfüllt wird, muß für $\beta_1 \geq 0$,

$$0 < 4(2\beta_2 - 3\beta_1 - 6)(4\beta_2 - 3\beta_1) < \beta_1(\beta_2 + 3)^2 \quad (3.39)$$

gelten. Daraus erkennt man, daß das zulässige Gebiet nach unten durch die Gerade gemäß Gl. (3.37b) und nach oben durch die Kurve nach Gl. (3.38b) begrenzt ist. Das Gebiet der Pearsonschen Kurven vom Typ VI liegt daher, wie in Bild 3.2 zu erkennen ist, zwischen den Gebieten der Typen I und IV.

Es sollen nun die drei Grundtypen des Pearsonschen Kurvensystems und die Spezialfälle im Detail behandelt werden.

4.2.1 Die Pearsonschen Kurven vom Typ I

Für $K < 0$ sind die Nullstellen von $B(z)$ reell und haben unterschiedliche Vorzeichen. Es sei z_1 die negative und z_2 die positive Nullstelle.

Zur Lösung der Differentialgleichung (3.29) wird das Polynom $B(z)$ zunächst in faktorisierter Form geschrieben

$$\frac{1}{p(z)} \frac{dp(z)}{dz} = \frac{z+a}{b_2(z-z_1)(z-z_2)} .$$

Die rechte Seite wird nun in die Partialbruchdarstellung

$$\frac{1}{p(z)} \frac{dp(z)}{dz} = \frac{z_1+a}{b_2(z_1-z_2)} \frac{1}{z-z_1} - \frac{z_2+a}{b_2(z_2-z_1)} \frac{1}{z_2-z} \quad (3.40)$$

gebracht. Die Integration dieser Differentialgleichung liefert mit

$$k_1 = \frac{z_1+a}{b_2(z_1-z_2)} \quad (3.41a)$$

und

$$k_2 = \frac{z_2+a}{b_2(z_2-z_1)} \quad (3.41b)$$

das Ergebnis

$$p(z) = c(z-z_1)^{k_1}(z_2-z)^{k_2} . \quad (3.42)$$

Der Ursprung des Koordinatensystems für Gl. (3.42) liegt im Mittelwert m_1 , da die Koeffizienten b_0 , b_1 , b_2 und a unter der Voraussetzung berechnet wurden, daß $m_1 = 0$ ist. Wird in Gl. (3.42) für z , $z-m_1$ eingesetzt, so wird das Koordinatensystem wieder auf den eigentlichen Ursprung von z versetzt. Man erhält dann

$$p(z) = c(z-\alpha)^{k_1}(\beta-z)^{k_2} . \quad (3.43)$$

Hierbei ist $\alpha = z_1 + m_1$ (3.44a)

und $\beta = z_2 + m_1$. (3.44b)

Da die Wahrscheinlichkeitsdichte $p(z)$ nicht negativ werden kann, ist der Variationsbereich von z auf das Intervall (α, β) beschränkt.

Die Integrationskonstante c in Gl. (3.43) wird aus der Forderung

$$\int_{\alpha}^{\beta} p(z) dz = 1$$

ermittelt. Daraus erhält man [43]

$$c = \frac{1}{(\beta - \alpha) k_1 + k_2 + 1} \cdot \frac{\Gamma(k_1 + k_2 + 2)}{\Gamma(k_1 + 1) \Gamma(k_2 + 1)}, \quad (3.45)$$

wobei, auf Grund der Definition der Gamma-Funktion, $k_1, k_2 > -1$ gelten muß. Daß diese Bedingungen für $K < 0$ immer erfüllt sind, kann durch Einsetzen der Gln. (3.33) - (3.35) in die Gln. (3.41) nach einigen algebraischen Manipulationen nachgewiesen werden.

Bild 3.3 zeigt die verschiedenen Formen, welche die Funktion $p(z)$ nach Gl. (3.43) in Abhängigkeit der Parameter k_1 und k_2 annehmen kann. Für alle diese Kurven wurde $\alpha = -1$ und $\beta = 1$ gewählt. Die geläufigste Form der Pearson-Kurve vom Typ I ist die "Glocken"-Form wie in Bild 3.3a. Sie tritt dann auf, wenn $k_1, k_2 > 0$ ist. Für $k_1 \leq 0$ und $k_2 \geq 0$, oder $k_1 \geq 0$ und $k_2 \leq 0$ hat die Kurve eine J-Form. Für $k_1, k_2 < 0$ nimmt sie dagegen eine U-Form an. Diese beiden Fälle werden in Bild 3.3b gezeigt. Selbstverständlich können die Kurven in Bild 3.3a und b auch eine zum gezeigten Verlauf symmetrische Form haben. Dazu brauchen lediglich die Werte von k_1 und k_2 vertauscht zu werden. Die Gebiete, die diesen drei Fällen im (β_1, β_2) -Diagramm entsprechen, sind im Bild 3.2 mit I, I(J) und I(U) gekennzeichnet.

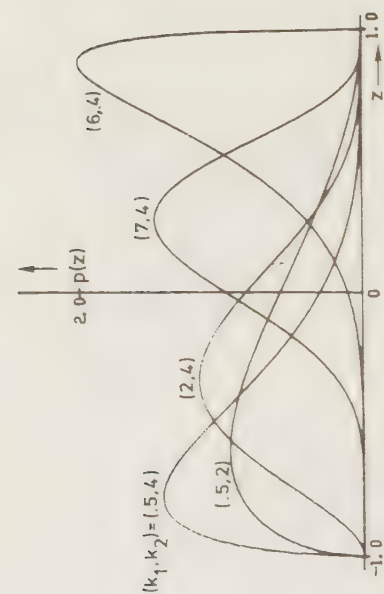
Die Grenze zwischen den Gebieten I und I(J) wird durch die Bedingungen $k_1 = 0$ und $k_2 > 0$ oder durch $k_2 = 0$ und $k_1 > 0$ definiert. Andererseits ist die Grenze zwischen den Gebieten I(J) und I(U) durch die Bedingungen $k_1 < 0$ und $k_2 = 0$ oder durch $k_2 < 0$ und $k_1 = 0$ gegeben. Die gestrichelte Kurve in Bild 3.2, die das Gebiet I(J) nach oben und unten abgrenzt, muß daher die Bedingung

$$k_1 \cdot k_2 \stackrel{!}{=} 0 \quad (3.46)$$

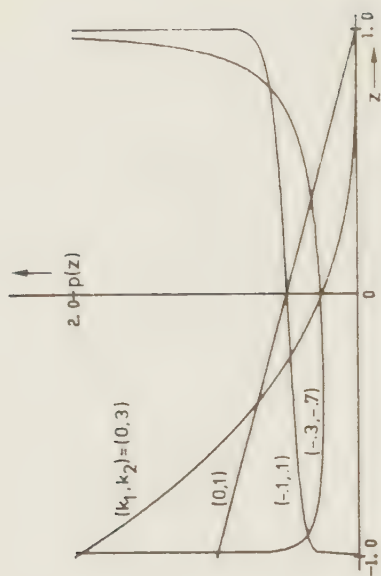
erfüllen. Mit k_1 und k_2 nach Gln. (3.41a,b) ist

$$k_1 \cdot k_2 = \frac{(z_1 + a)(z_2 + a)}{b_2^2 (z_1 - z_2)(z_2 - z_1)} \stackrel{!}{=} 0$$

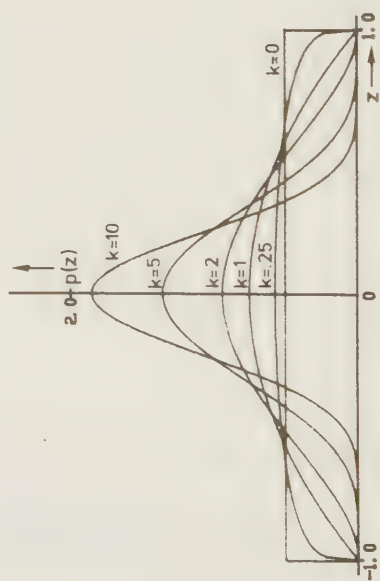
Wird hier z_1 und z_2 nach Gl. (3.34) eingesetzt, ergibt sich nach



(a) Typ I: $k_1, k_2 = 0$



(b) Typ I (J): $k_1 \leq 0, k_2 \geq 0$; oder $k_2 = 0, k_1 < 0$
Typ I (U): $-1 < k_1, k_2 < 0$



(c) Typ II: $k = k_1 = k_2 \geq 0$



(d) Typ II (U): $-1 < k_1 = k_2 < 0$

Bild 3.3: Verschiedene Formen der Pearsonschen Kurven vom Typ I und II

einigen Zwischenrechnungen folgende Gleichung für die Grenze des I(J)-Gebietes:

$$4(4\beta_2 - 3\beta_1)(5\beta_2 - 6\beta_1 - 9)^2 - \beta_1(\beta_2 + 3)^2(8\beta_2 - 9\beta_1 - 12) = 0. \quad (3.47)$$

Ein Sonderfall der Kurven vom Typ I(J) liegt vor, wenn das Wertepaar (β_1, β_2) auf die Gerade $5\beta_2 - 6\beta_1 - 9 = 0$ fällt (vgl. Bild 3.2). Es kann nachgewiesen werden, daß in diesem Fall $k_1 = -k_2$ ist. Die Kurven haben dann die Form eines verbogenen J wie in Bild 3.3b für $k_1 = -0,1$ und $k_2 = 0,1$ gezeigt wird. Die Gln. (3.34) und (3.41a,b) können jetzt nicht mehr für die Berechnung von z_1, z_2 bzw. k_1, k_2 herangezogen werden, da in den Gln. (3.33a-d) der Nenner verschwindet. Durch Einführen der Gln. (3.33a-d) in die Gln. (3.34) und (3.41a,b) erhält man jedoch, mit β_1 und β_2 nach Gln. (3.31) und (3.32), für den Fall, daß $5\beta_2 - 6\beta_1 - 9 = 0$ ist,

$$z_{1,2} = \frac{1}{\mu_2} (\mu_3 \mp \sqrt{\mu_3^2 + 3\mu_2^3}) \quad (3.48)$$

und

$$k_{1,2} = \mp \frac{\mu_3}{\sqrt{\mu_3^2 + 3\mu_2^3}}, \quad (3.49)$$

wobei das Minuszeichen jeweils für z_1 bzw. k_1 zu verwenden ist.

Es soll nun die absolut unterste Grenze für die (β_1, β_2) -Werte ermittelt werden. Aus der o.g. Bedingung $k_1, k_2 > -1$ folgt, daß der Extremfall bei $k_1 = k_2 = -1$ liegt. Die Bedingungsgleichung für die unterste Grenze ist daher

$$k_1 + k_2 \stackrel{!}{=} -2 \quad (3.50)$$

Mit k_1 und k_2 nach Gln. (3.41a,b) ist

$$k_1 + k_2 = \frac{1}{b_2} \stackrel{!}{=} -2$$

und mit b_2 gemäß Gl. (3.33d) folgt daraus die Gerade

$$\beta_2 - \beta_1 - 1 = 0 \quad (3.51)$$

als unterste Grenze für die (β_1, β_2) -Werte.

In den Bildern 3.3c,d werden Kurven vom Typ II gezeigt. Diese Kurven sind symmetrisch, da $k_1 = k_2$ ist. Wie in Abschnitt 4.2.4 noch gezeigt wird, können die Kurven vom Typ II als Sonderfall der Kurven vom Typ I betrachtet werden.

Die Pearsonschen Kurven vom Typ I mit Wahrscheinlichkeitsdichtefunktionen $p(z)$ gemäß Gl. (3.42) oder (3.43) entsprechen den in der Literatur üblicherweise mit Beta-Verteilung bezeichneten Funktionen. Die Beta-Verteilung ist daher nichts anderes als ein Sonderfall des Pearsonschen Kurvensystems.

4.2.2 Die Pearsonschen Kurven vom Typ IV

Wenn $0 < K < 1$ ist, erhält man zwei komplex konjugierte Nullstellen. Die rechte Seite der Differentialgleichung (3.29) läßt sich nun nicht mehr in die Partialbruchdarstellung Gl. (3.40) mit reellen Faktoren bringen. Mit der Koordinatentransformation

$$\bar{z} = z + \frac{b_1}{2b_2} \quad (3.52)$$

wird aus Gl. (3.29)

$$\frac{d}{d\bar{z}} \left(\ln p(\bar{z}) \right) = \frac{\bar{z} + \bar{a}}{b_2(\bar{z}^2 + d^2)} \quad (3.53)$$

Hierbei ist

$$\bar{a} = a - \frac{b_1}{2b_2} \quad (3.54a)$$

und

$$d^2 = \frac{b_0}{b_2} - \frac{b_1^2}{4b_2^2} \quad (3.54b)$$

Die Lösung der Differentialgleichung (3.52) ergibt

$$\ln p(\bar{z}) = \frac{1}{2b_2} \ln (\bar{z}^2 + d^2) + \frac{\bar{a}}{b_2 d} \arctan \frac{\bar{z}}{d} + c'$$

oder

$$p(\bar{z}) = c(1 + \bar{z}^2/d^2)^{-m} e^{-v \arctan \frac{\bar{z}}{d}} \quad (3.55)$$

worin

$$m = - \frac{1}{2b_2} \quad (3.56a)$$

und

$$v = - \frac{\bar{a}}{b_2 d} \quad (3.56b)$$

ist. In Gl. (3.55) ist darauf zu achten, daß der Mittelwert nicht im Ursprung, sondern bei $\bar{z} = b_1/2b_2$ liegt. Mit der Koordinaten-Rücktransformation nach Gl. (3.52) erhält man

$$p(z) = c \left(1 + \left(\frac{z}{d} - \frac{v}{r} \right)^2 \right)^{-m} e^{-v \arctan \left(\frac{z}{d} - \frac{v}{r} \right)} \quad (3.57)$$

mit

$$r = 2m - 2 \quad (3.58)$$

Der Variationsbereich von $\bar{z}$ bzw. z erstreckt sich von $-\infty$ bis $+\infty$.

Die Bestimmung der Integrationskonstante ist hier nicht ganz so einfach, wie im Fall der Kurven vom Typ I, da keine geschlossene Lösung für c aus der Bedingungsgleichung

$$\int_{-\infty}^{\infty} c \left(1 + \bar{z}^2/d^2 \right)^{-m} e^{-v \arctan \frac{\bar{z}}{d}} d\bar{z} = 1 \quad (3.59)$$

angegeben werden kann. Dieses Integral kann nur numerisch berechnet werden. Dazu ist es erforderlich, durch eine geeignete Variablentransformation die Integrationsgrenzen auf ein endliches Intervall zu begrenzen. Mit der Transformation

$$\phi = \arctan \left(\frac{z}{d} \right) \quad (3.60)$$

kann dies erreicht werden, da sich dann Gl. (3.59) vereinfacht auf

$$\frac{1}{c} = d \int_{-\pi/2}^{\pi/2} \cos^r \phi e^{-v\phi} d\phi \quad (3.61)$$

Dieses Integral läßt sich nun auf einfache Weise, z.B. nach Simpson, auf dem Digitalrechner numerisch berechnen.

Im Bild 3.4 wird an Hand einiger Beispiele gezeigt, welche For-

men die Pearsonschen Kurven vom Typ IV in Abhängigkeit von Schiefe und Exzeß haben können. Für alle Kurven gilt $m_1 = 0$ und $\mu_2 = 1$. Kurven, deren Maximum bei positiven Werten von z liegt, haben ein negatives μ_3 und umgekehrt. Im Bild 3.4 wird eine symmetrische Kurve mit $\beta_1 = 0$ gezeigt. Dieser Sonderfall des Kurventyps IV wird im Abschnitt 4.2.5 noch ausführlich behandelt.

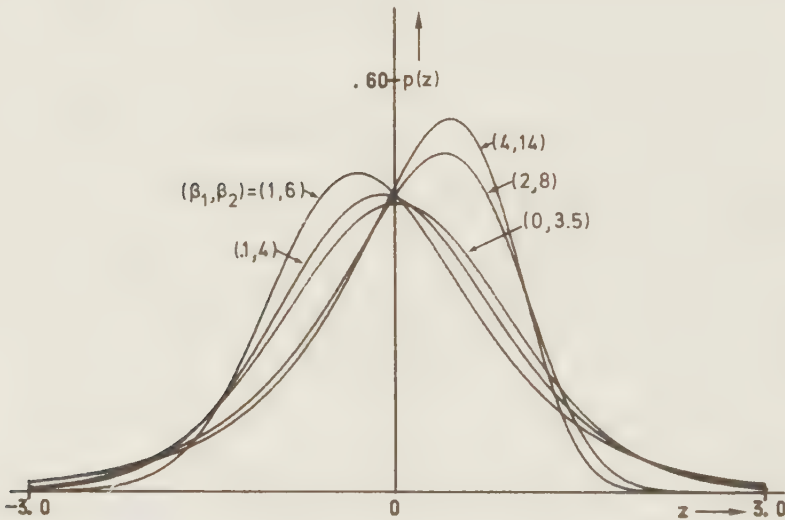


Bild 3.4: Verschiedene Formen der Pearsonschen Kurven vom Typ IV ($m_1 = 0$, $\mu_2 = 1$)

4.2.3 Die Pearsonschen Kurven vom Typ VI

Für den Fall, daß $K > 1$ ist, erhält man wegen $b_1^2 > 4b_0b_2$ in Gl. (3.34) zwei reelle Nullstellen mit gleichem Vorzeichen

$$z_1 = -\frac{b_1}{2b_2} + \frac{1}{2b_2} \sqrt{b_1^2 - 4b_0b_2} \quad (3.62a)$$

und

$$z_2 = -\frac{b_1}{2b_2} - \frac{1}{2b_2} \sqrt{b_1^2 - 4b_0b_2} . \quad (3.62b)$$

Im Bereich der Pearsonschen Kurven vom Typ VI gilt, wie aus dem (β_1, β_2) -Diagramm in Bild 3.2 sofort zu erkennen ist,

$$2\beta_2 - 3\beta_1 - 6 > 0 , \quad (3.63a)$$

$$5\beta_2 - 6\beta_1 - 9 > 0 , \quad (3.63b)$$

$$4\beta_2 - 3\beta_1 > 0 . \quad (3.63c)$$

Damit ist gemäß Gln. (3.33a-e)

$$a = \text{sign}(\mu_3) |a| ,$$

$$b_0 < 0 , \quad (3.64)$$

$$b_1 = -a ,$$

$$b_2 < 0 .$$

Unter Berücksichtigung dieser Bedingungen können, je nachdem, ob $\mu_3 > 0$ oder $\mu_3 < 0$ ist, folgende zwei Fälle für die Lage der Nullstellen von $B(z)$ gemäß Gln. (3.62a,b) unterschieden werden:

a) $\mu_3 > 0$

Für $\mu_3 > 0$ ist $b_1 < 0$. Die beiden Nullstellen sind negativ und befinden sich im Bereich

$$-\left|\frac{b_1}{b_2}\right| < z_1 < -\left|\frac{b_1}{2b_2}\right|$$

$$-\left|\frac{b_1}{2b_2}\right| < z_2 < 0 .$$

b) $\mu_3 < 0$

Für $\mu_3 < 0$ ist $b_1 > 0$. Die beiden Nullstellen sind positiv und

befinden sich im Bereich

$$0 < z_1 < \left| \frac{b_1}{2b_2} \right|$$

$$\left| \frac{b_1}{2b_2} \right| < z_2 < \left| \frac{b_1}{2b_2} \right|$$

In beiden Fällen gilt also $z_1 < z_2$.

Da es sich hier um reelle Nullstellen handelt, kann die Differentialgleichung (3.29) in ähnlicher Weise wie für den Typ I gelöst werden. Da voraussetzungsgemäß der Mittelwert mit dem Ursprung des Koordinatensystems zusammenfällt, kann jedoch in diesem Fall der Variationsbereich von z nicht, wie für die Kurven vom Typ I, mit dem Intervall (z_1, z_2) übereinstimmen. Die beiden Nullstellen liegen hier nämlich entweder gleichzeitig rechts oder links vom Ursprung. Der Variationsbereich von z muß daher außerhalb dieses Intervalls liegen und zwar so, daß er den Ursprung des Koordinatensystems (wo der Mittelwert von $p(z)$ liegt) mit einschließt. Aus diesem Grund müssen beide Fälle getrennt behandelt werden.

a) $u_2 > 0$.

Beide Nullstellen sind negativ ($z_1 < z_2 < 0$). $p(z)$ muß daher für $z > z_2$ definiert sein (Bild 3.5a). Aus diesem Grund wird die rechte Seite der Differentialgleichung (3.29) in die Partialbruchdarstellung

$$\frac{1}{p(z)} \frac{d p(z)}{dz} = \frac{z_1 + a}{b_2(z_1 - z_2)} \frac{1}{z - z_1} + \frac{z_2 + a}{b_2(z_2 - z_1)} \frac{1}{z - z_2} \quad (3.65)$$

gebracht. Daraus ergibt sich

$$p(z) = c(z - z_1)^{k_1} (z - z_2)^{k_2}, \quad z > z_2 \quad (3.66)$$

mit k_1 und k_2 gemäß Gln. (3.41a,b).

Die Integrationskonstante c ist aus der Forderung

$$c \int_{z_2}^{\infty} (z-z_1)^{k_1} (z-z_2)^{k_2} dz = 1 \quad (3.67)$$

zu bestimmen. Mit der Variablenformation

$$x = \frac{z-z_2}{z_2-z_1},$$

wird aus Gl. (3.67)

$$c(z_2-z_1)^{k_1+k_2+1} \int_0^{\infty} (x+1)^{k_1} x^{k_2} dx = 1. \quad (3.68)$$

Für das vorkommende bestimmte Integral läßt sich eine geschlossene Lösung angeben [43], mit der dann c aus Gl. (3.68) berechnet werden kann

$$c = (z_2-z_1)^{-k_1-k_2-1} \frac{\Gamma(-k_1)}{\Gamma(k_2+1) \Gamma(-k_1-k_2-1)}. \quad (3.69)$$

Hierbei muß $-k_1 > k_2 + 1 > 0$ sein. Daß diese Bedingungen für die Pearsonschen Kurven vom Typ VI erfüllt werden, soll im folgenden bewiesen werden.

Mit k_2 aus Gl. (3.41b) und z_1, z_2 aus Gln. (3.62a,b) ist

$$k_2+1 = \frac{-z_2(1+2b_2)}{\sqrt{b_1^2 - 4b_0b_2}}.$$

Da für die Kurven vom Typ VI $K > 1$ gilt, ist $b_1^2 - 4b_0b_2 > 0$. Für $\mu_3 > 0$ ist außerdem $z_2 < 0$. Daher ist die Bedingung $k_2+1 > 0$ erfüllt, wenn

$$1 + 2b_2 > 0$$

ist. Wird hier b_2 aus Gl. (3.33d) eingeführt, entsteht daraus die Forderung

$$\frac{2-\beta_1-1}{5\beta_2-6\beta_1-9} > 0.$$

Aus der Betrachtung des (β_1, β_2) -Diagramms in Bild 3.2 kann sofort entnommen werden, daß diese Bedingung im Gebiet der Pear-

sonischen Kurven vom Typ VI immer erfüllt ist. Damit wurde bewiesen, daß $k_2 + 1 > 0$ ist. Es fehlt noch der Beweis, daß $-k_1 > k_2 + 1$ ist, oder mit k_1 und k_2 aus Gln. (3.41a,b), daß

$$-\frac{z_1 + a}{b_2(z_1 - z_2)} > \frac{z_2 + a}{b_2(z_2 - z_1)} + 1$$

ist. Dies führt unter Berücksichtigung der Ungleichungen $b_2 < 0$ und $z_1 < z_2 < 0$ letztlich zu der Forderung

$$1 + b_2 > 0$$

und mit b_2 aus Gl. (3.33d) zu

$$\frac{8\beta_2 - 9\beta_1 - 12}{5\beta_2 - 6\beta_1 - 9} > 0.$$

Wird die Gerade $8\beta_2 - 9\beta_1 - 12 = 0$ in das (β_1, β_2) -Diagramm (Bild 3.2) eingezeichnet, kann nachgeprüft werden, daß sie im gesamten Bereich $\beta_1 \geq 0$ unterhalb der unteren Grenze des Gebietes VI liegt. In diesem Gebiet gilt also $8\beta_2 - 9\beta_1 - 12 > 0$ und auf Grund von Gl. (3.63c) $5\beta_2 - 6\beta_1 - 9 > 0$. Die Bedingung $-k_1 > k_2 + 1$ wird daher im Gebiet VI immer erfüllt.

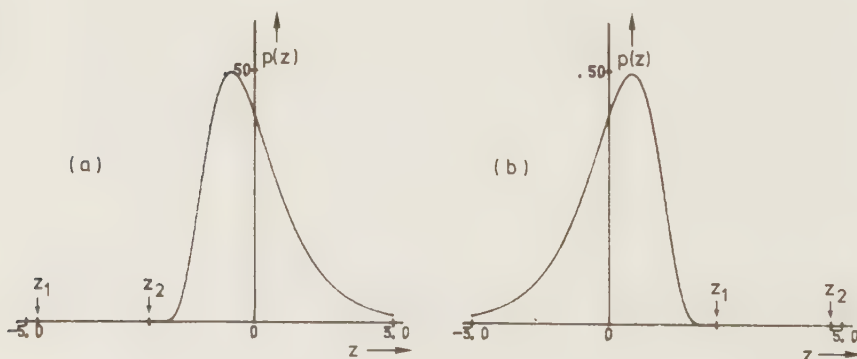


Bild 3.5: Pearsonsche Kurven vom Typ VI mit Lage der Nullstellen des Polynoms $B(z)$ für $\mu_3 > 0$ (a) und $\mu_3 < 0$ (b).

Im Bild 3.6 wird an Hand einiger Beispiele gezeigt, welche Formen die Pearsonschen Kurven vom Typ VI für $\mu_3 > 0$ in Abhängigkeit von Schiefe und Exzeß haben können. Im allgemeinen werden sie glockenförmig sein, für $k_2 \leq 0$ sind sie jedoch J-förmig. Dies ist dann der Fall, wenn sich das (β_1, β_2) -Wertepaar im Gebiet VI genau auf oder unter der gestrichelten Kurve in Bild 3.2 befindet. Diese Kurve wurde bereits im Abschnitt 4.2.1 berechnet und ist durch Gl. (3.47) gegeben.

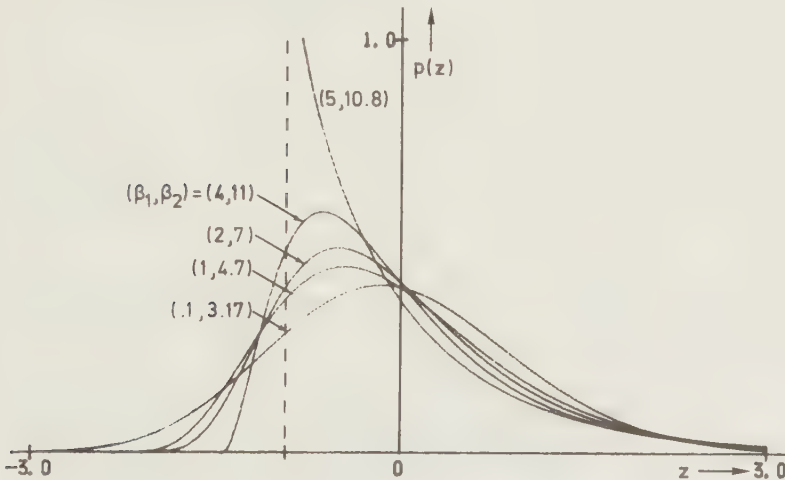


Bild 3.6: Verschiedene Formen der Pearsonschen Kurven vom Typ VI für $\mu_3 > 0$ ($m_1 = 0$, $\mu_2 = 1$).

b) $\mu_3 < 0$

Beide Nullstellen sind positiv ($0 < z_1 < z_2$). $p(z)$ muß daher für $z < z_1$ definiert sein (Bild 3.5b). Aus diesem Grund wählt man an Stelle der Partialbruchdarstellung von Gl. (3.65) die Zerlegung

$$\frac{1}{p(z)} \frac{dp(z)}{dz} = - \frac{z_1 + a}{b_2(z_1 - z_2)} \frac{1}{z_1 - z} - \frac{z_2 + a}{b_2(z_2 - z_1)} \frac{1}{z_2 - z} . \quad (3.70)$$

Daraus ergibt sich die Lösung

$$p(z) = c(z_1 - z)^{k_1} (z_2 - z)^{k_2}, \quad z < z_1 \quad (3.71)$$

mit k_1 und k_2 gemäß Gln. (3.41a,b).

Die Berechnung der Integrationskonstante erfolgt hier aus der Forderung

$$c \int_{-\infty}^{z_1} (z_1 - z)^{k_1} (z_2 - z)^{k_2} dz = 1.$$

Ähnlich wie für den Fall $\mu_3 > 0$ erhält man daraus

$$c = (z_2 - z_1)^{-k_1 - k_2 - 1} \frac{\Gamma(-k_2)}{\Gamma(k_1 + 1) \Gamma(-k_1 - k_2 - 1)} \quad (3.72)$$

wobei $-k_2 > k_1 + 1 > 0$ gelten muß. Daß diese Bedingungen erfüllt sind, kann in völlig analoger Weise zum Fall $\mu_3 > 0$ bewiesen werden.

Die Kurven vom Typ VI sind für $\mu_3 < 0$ symmetrisch zu den Kurven für $\mu_3 > 0$ bei gleichen Werten von β_1 und β_2 (vgl. Bild 3.6).

Außer den bisher behandelten drei Grundtypen der Pearsonschen Kurven gibt es einige Sondertypen, die im folgenden aufgeführt werden. Es handelt sich dabei um Spezialfälle, die auf dem Rand der Gebiete der Grundtypen oder im Übergangsbereich zwischen ihnen auftreten.

4.2.4 Die Pearsonschen Kurven vom Typ II

Für $K = 0$ ist nach Gl. (3.35) $\beta_1 = 0$. Im (β_1, β_2) -Diagramm befindet man sich daher auf der β_2 -Achse. Ist außerdem $\beta_2 < 3$, kann dieser Kurventyp als Sonderfall des Typs I betrachtet werden, da er genau auf dem Rand des Gebietes I auftritt (Bild 3.2).

Mit $\beta_1 = 0$ und $\beta_2 < 3$ ist nach Gln. (3.31) und (3.33a-e)

$$a = b_1 = 0 \quad (3.73a)$$

und
$$\frac{b_0}{b_2} < 0. \quad (3.73b)$$

Die beiden Nullstellen nach Gl. (3.34) sind daher reellwertig und dem Betrage nach gleich

$$z_0 = -z_1 = z_2 = \sqrt{-b_0/b_2}. \quad (3.74)$$

Mit Gln. (3.73a) und (3.74) in Gln. (3.41a,b) ist

$$k = k_1 = k_2 = \frac{1}{2b_2}. \quad (3.75)$$

Mit den Gln. (3.74) und (3.75) erhält man aus der Lösung der Differentialgleichung nach Gl. (3.42) die symmetrische Kurve

$$p(z) = c(z_0^2 - z^2)^k. \quad (3.76)$$

Der Variationsbereich von z liegt im Intervall $(-z_0, z_0)$ und die Integrationskonstante wird nach Gl. (3.45) berechnet. Diese vereinfacht sich auf Grund der Gln. (3.74) und (3.75) zu

$$c = \frac{1}{(2z_0)^{2k+1}} \frac{\Gamma(2k+2)}{(\Gamma(k+1))^2} \quad (3.77)$$

Die Kurven nach Gl. (3.76) haben für $k > 0$ Glockenform (Bild 3.3c) und für $k < 0$ U-Form (Bild 3.3d). Für $k = 0$ erhält man die Gleichverteilung (Bild 3.3c). Diese tritt, wie mit Hilfe der Gln. (3.75) und (3.33d,e) nachgeprüft werden kann, für $\beta_1 = 0$ und $\beta_2 = 1,8$ auf.

4.2.5 Die Pearsonschen Kurven vom Typ VII

Wie für den Kurventyp II gilt hier $K = 0$ und $\beta_1 = 0$. Der Unterschied zum Typ II besteht darin, daß hier statt $\beta_2 < 3$ nun

$\beta_2 > 3$ ist. Wegen $\beta_1 = 0$ handelt es sich hier ebenfalls um eine symmetrische Verteilung, die, da $\beta_2 > 3$ ist, "spitzer" als die Normalverteilung ist.

Mit $\beta_1 = 0$ und $\beta_2 > 3$ ist nach Gln. (1.31) und (3.33a-e)

$$a = b_1 = 0 \quad (3.78a)$$

und
$$\frac{b_0}{b_2} > 0. \quad (3.78b)$$

Die beiden Nullstellen nach Gl. (3.34) sind daher rein imaginär und zueinander konjugiert. Dieser Kurventyp kann also als Sonderfall des Typs IV betrachtet werden. Mit Gl. (3.78a) in Gln. (3.54a) und (3.56b) ist $v = 0$ und Gl. (3.55) vereinfacht sich zu

$$p(z) = c(1+z^2/d^2)^{-m}. \quad (3.79)$$

Wegen $d^2 = b_0/b_2 > 0$ ist hier $p(z)$, wie für die Kurven vom Typ IV, im gesamten Bereich $-\infty < z < \infty$ definiert. Die Integrationskonstante c wird aus Gl. (3.61) mit $v = 0$ berechnet, d.h. aus

$$\frac{1}{c} = d \int_{-\pi/2}^{\pi/2} \cos^r \phi \, d\phi \quad (3.80)$$

Dieses Integral läßt sich nun im Gegensatz zu dem in Gl. (3.61) geschlossen lösen [43]. Es ist

$$\frac{1}{c} = 2d \frac{\Gamma(\frac{r+1}{2}) \Gamma(\frac{1}{2})}{2\Gamma(\frac{r+2}{2})} \quad (3.81)$$

wobei $r > -1$ sein muß. Mit Hilfe der Gln. (3.58), (3.56a) und (3.33d,e) kann leicht nachgewiesen werden, daß für $\beta_1 = 0$, $\beta_2 > 3$ diese Bedingung immer erfüllt ist. Mit $\Gamma(\frac{1}{2}) = \pi$ und $r = 2m-2$ (Gl. (3.58)) ist schließlich

$$c = \frac{\Gamma(m)}{d\sqrt{\pi} \Gamma(m-\frac{1}{2})}. \quad (3.82)$$

Im Bild 3.7 werden einige Beispiele der Pearsonschen Kurven vom Typ VII für gleichbleibenden Mittelwert und Varianz ($m_1=0$, $\mu_2=1$) und verschiedene Werte von β_2 gezeigt.

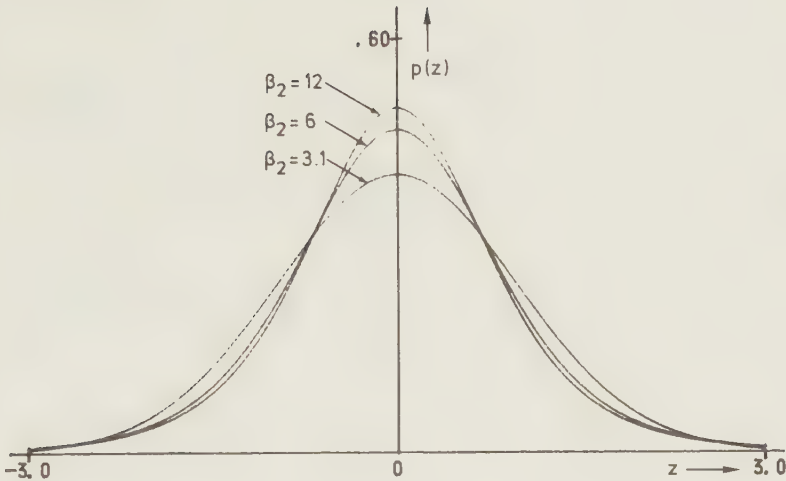


Bild 3.7: Verschiedene Formen der Pearsonschen Kurven vom Typ VII ($m_1 = \beta_1 = 0$, $\mu_2 = 1$)

4.2.6 Die Pearsonschen Kurven vom Typ III

Der Kurventyp III tritt dann auf, wenn sich die Wertepaare (β_1, β_2) genau auf der Grenze zwischen den Gebieten I und VI, also auf der Geraden $2\beta_2 - 3\beta_1 - 6 = 0$, befinden (Bild 3.2). Dann gilt gemäß Gl. (3.33d) $b_2 = 0$ und gemäß Gl. (3.35) $K \rightarrow \infty$. Mit $b_2 = 0$ reduziert sich die Differentialgleichung (3.29) auf

$$\frac{d(\ln p(z))}{dz} = \frac{z+a}{b_0+b_1 z} \quad (3.83)$$

Durch Integration erhält man daraus nach einigen Zwischenschritten

$$p(z) = c e^{-\gamma z} \left(1 + \frac{z}{d}\right)^k \quad (3.84)$$

mit

$$d = \frac{b_0}{b_1}, \quad (3.85a)$$

$$\gamma = -\frac{1}{b_1}, \quad (3.85b)$$

$$k = -\frac{b_1^{2+b_0}}{b_1^2}. \quad (3.85c)$$

Wird in Gl. (3.85a) b_0 und b_1 nach Gln. (3.33b,c) eingeführt, dann erhält man für d , auf der Geraden $2\beta_2 - 3\beta_1 - 6 = 0$

$$d = \frac{2\mu_2^2}{\mu_3}. \quad (3.86)$$

d hat also das gleiche Vorzeichen wie μ_3 .

Der Variationsbereich von z in Gl. (3.84) ist daher wie folgt definiert (Bild 3.8):

$$\mu_3 > 0: \quad -d < z < \infty$$

$$\mu_3 < 0: \quad -\infty < z < -d$$

und der Ursprung des Koordinatensystems liegt voraussetzungs-
gemäß im Mittelwert m_1 .

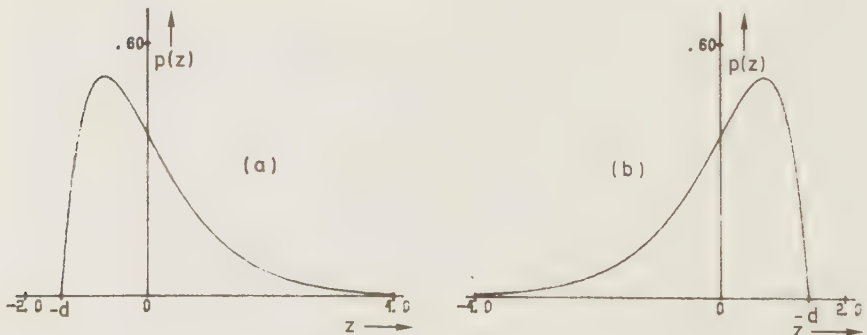


Bild 3.8: Pearsonsche Kurven vom Typ III für $\mu_3 > 0$ (a)
und $\mu_3 < 0$ (b)

Die Integrationskonstante c in Gl. (3.84) wird, für den Fall, daß $\mu_3 > 0$ ist, aus der Forderung

$$\int_{-d}^{\infty} p(z) dz = 1 \quad (3.87)$$

berechnet. Dies führt zu dem Ergebnis

$$c = \frac{\gamma (k+1)^k}{e^{k+1} \Gamma(k+1)} . \quad (3.88)$$

Für den Fall, daß $\mu_3 < 0$ ist, wird in Gl. (3.87) über das Intervall $(-\infty, -d)$ integriert. Daraus folgt

$$c = \frac{-\gamma (k+1)^k}{e^{k+1} \Gamma(k+1)} . \quad (3.89)$$

Da auf Grund von Gl. (3.85b) und Gln. (3.33a,c)

$$\gamma = \text{sign}(\mu_3) \left| \frac{1}{a} \right|$$

ist, können die Gln. (3.88) und (3.89) zusammengefaßt werden in

$$c = \frac{|\gamma| (k+1)^k}{e^{k+1} \Gamma(k+1)} . \quad (3.90)$$

In den Gln. (3.88-90) muß $k > -1$ gelten. Ähnlich wie in den früheren Fällen, kann auch hier nachgewiesen werden, daß diese Bedingung für die Kurven vom Typ III immer erfüllt wird.

Im Bild 3.9 werden einige Beispiele der Pearsonschen Kurven vom Typ III für $\mu_3 > 0$ und verschiedene Werte von β_1 und β_2 gezeigt. In allen Fällen gilt $m_1 = 0$ und $\mu_2 = 1$. Die Kurven für $\mu_3 < 0$ sind symmetrisch zu den entsprechenden Kurven für $\mu_3 > 0$.

Aus Gl. (3.84) ist ersichtlich, daß die Kurven für $k > 0$ glockenförmig und für $k \leq 0$ J-förmig sind. Auf Grund der Gl. (3.85c) und der Gln. (3.33a,b,c) bedeutet dies, daß für (β_1, β_2) -Wertepaare, die auf der Geraden $2\beta_2 - 3\beta_1 - 6 = 0$ unterhalb des Punktes

$\beta_1 = 4, \beta_2 = 9$ liegen (Bild 3.2), die Kurven glockenförmig sind und für Werte, die oberhalb dieses Punktes liegen, die Kurven J-förmig sind. Im Punkt $\beta_1 = 4, \beta_2 = 9$ ist $k = 0$ und die Gl. (3.84) geht in die Gleichung für die Exponentialverteilung

$$p(z) = c e^{-\gamma z} \quad (3.91)$$

über.

Eine Verteilung, die durch Gl. (3.84) beschrieben wird, wird üblicherweise Gammaverteilung genannt. Die Gamma- und die Exponentialverteilung können daher ebenso wie die Betaverteilung als Sonderfälle des Pearsonschen Kurvensystems betrachtet werden.

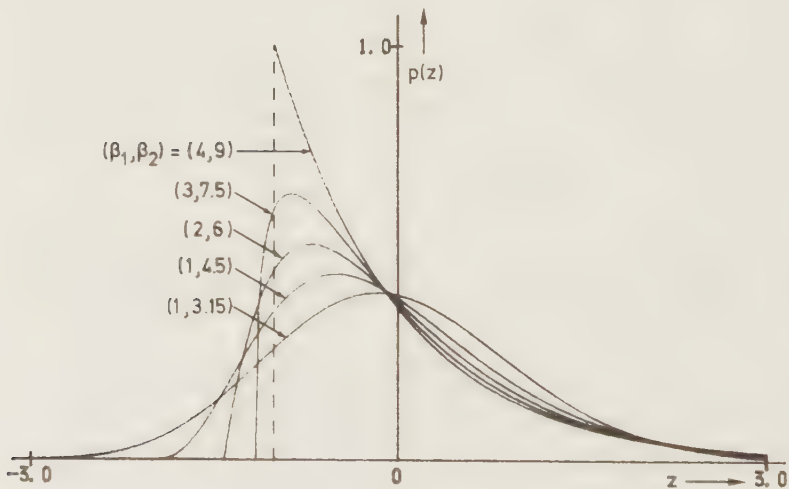


Bild 3.9: Verschiedene Formen der Pearsonschen Kurven vom Typ III für $\mu_3 > 0$ ($\mu_1 = 0, \mu_2 = 1$)

4.2.7 Die Pearsonschen Kurven vom Typ V

Auf der Grenze zwischen den Gebieten der Kurven vom Typ IV und VI mit den Bedingungen $0 < K < 1$ bzw. $K > 1$, befindet sich der Kurventyp V mit der Bedingung $K = 1$. Gemäß Gl. (3.36) tritt dieser Fall dann auf, wenn sich das Wertepaar (β_1, β_2) genau auf der Kurve

$$4(2\beta_2 - 3\beta_1 - 6)(4\beta_2 - 3\beta_1) = \beta_1(\beta_2 + 3)^2 \quad (3.92)$$

befindet (Bild 3.2). Für $K = 1$ besitzt das Polynom $B(z)$ nach Gl. (3.34) und (3.35) die doppelt reelle Nullstelle

$$z_0 = -\frac{b_1}{2b_2}, \quad (3.93)$$

womit die Differentialgleichung (3.29) die Spezialform

$$\frac{d(\ln p(z))}{dz} = \frac{z+a}{b_2(z-z_0)^2} \quad (3.94)$$

hat. Werden in Gl. (3.93) b_1 und b_2 nach Gln. (3.33c,d) eingeführt, so ist

$$z_0 = -\frac{\mu_3(\beta_2 + 3)}{\mu_2(2\beta_2 - 3\beta_1 - 6)}.$$

Diese Beziehung besagt, daß

$$z_0 > 0 \quad \text{für} \quad \mu_3 < 0$$

$$\text{und} \quad z_0 < 0 \quad \text{für} \quad \mu_3 > 0$$

ist. Auf Grund ähnlicher Überlegungen wie im Abschnitt 4.2.3 für die Kurven vom Typ VI bedeutet dies, daß sich der Variationsbereich von z in Gl. (3.94) für den Fall $\mu_3 > 0$ von einem negativen Wert z_0 bis $+\infty$ und für den Fall $\mu_3 < 0$ von $-\infty$ bis zu einem positiven Wert z_0 erstreckt (Bild 3.10).

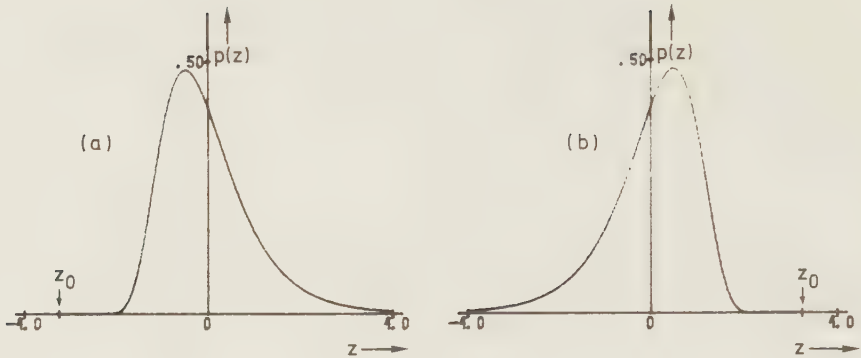


Bild 3.10: Pearsonsche Kurven vom Typ V für $\mu_3 > 0$ (a)
und $\mu_3 < 0$ (b)

Aus diesem Grund müssen auch hier beide Fälle getrennt behandelt werden.

a) $\mu_3 > 0$

Die Lösung der Differentialgleichung (3.94) liefert

$$p(z) = c(z-z_0)^{-k} e^{-\frac{\gamma}{z-z_0}}, \quad z > z_0 \quad (3.95)$$

mit

$$k = -\frac{1}{b_2} \quad (3.96a)$$

und
$$\gamma = \frac{z_0 + a}{b_2} \quad (3.96b)$$

Die Integrationskonstante c wird aus der Forderung

$$\int_{z_0}^{\infty} p(z) dz = 1$$

mit $p(z)$ aus Gl. (3.95) bestimmt. Man erhält daraus

$$c = \frac{\gamma^{k-1}}{\Gamma(k-1)}, \quad (3.97)$$

wobei $\gamma > 0$ und $k > 1$ sein muß.

b) $\mu_3 < 0$

In diesem Fall wird folgende Lösung der Differentialgleichung (3.94) verwendet:

$$p(z) = c(z_0 - z)^{-k} e^{\frac{\gamma}{z_0 - z}}, \quad z < z_0 \quad (3.98)$$

mit k und γ nach Gln. (3.96a,b). Die Integrationskonstante ergibt sich aus der Forderung

$$\int_{-\infty}^{z_0} p(z) dz = 1$$

und ist

$$c = \frac{(-\gamma)^{k-1}}{\Gamma(k-1)}, \quad (3.99)$$

wobei $\gamma < 0$ und $k > 1$ sein muß.

Wie in den vorausgegangenen Fällen kann auch hier nachgewiesen werden, daß die Bedingungen für γ und k in den Gln. (3.97) und (3.99) für die beiden Fälle $\mu_3 > 0$ und $\mu_3 < 0$ immer erfüllt werden.

In Bild 3.11 wird an Hand einiger Beispiele gezeigt, welche Formen die Pearsonschen Kurven vom Typ V für $\mu_3 < 0$ in Abhängigkeit von Schiefe und Exzeß haben können. Für alle Kurven gilt $m_1 = 0$ und $m_2 = 1$. Die Kurven für $\mu_3 > 0$ sind symmetrisch zu den entsprechenden Kurven für $\mu_3 < 0$.

Dieser Kurventyp entspricht den üblicherweise mit "Logarithmus-Normalverteilung" bezeichneten Kurven. Die Log.-Normalverteilung kann daher als Sonderfall des Pearsonschen Kurvensystems betrachtet werden.

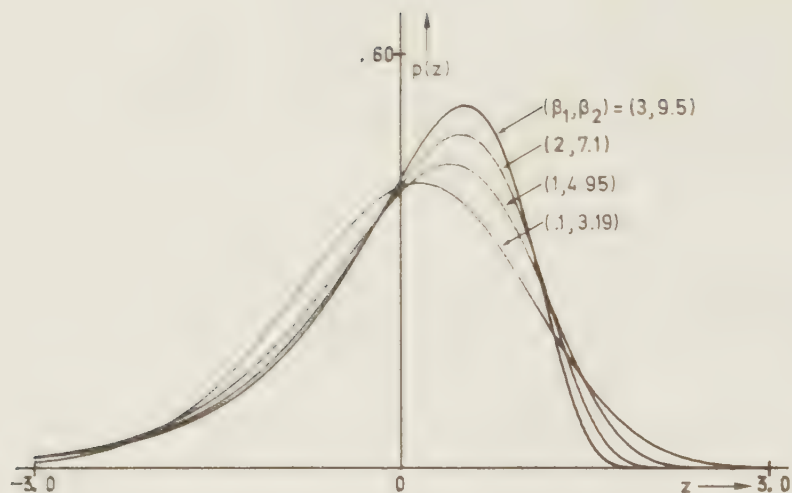


Bild 3.11: Verschiedene Formen der Pearsonischen Kurven vom Typ V für $\mu_3 < 0$ ($m_1 = 0$, $\mu_2 = 1$)

4.1.6 Die Normalverteilung

Es soll hier nur noch gezeigt werden, daß die Normalverteilung auch als Sonderfall des Pearsonischen Kurvensystems betrachtet werden kann.

Für die Normalverteilung ist $\beta_1 = 0$ und $\beta_2 = 3$ (siehe Anhang, Abschnitt 1.8). Im (s_1, s_2) -Diagramm (Bild 3.2) wurde dieser Punkt mit X gekennzeichnet.

Für $s_1 = 0$, $s_2 = 3$ gilt gemäß Gl.n. (3.33a-e)

$$a = b_1 = b_2 = 0$$

und $b_3 = -\frac{1}{2}$.

Damit vereinfacht sich die Differentialgleichung (3.29) auf

$$\frac{d\{\ln p(z)\}}{dz} = -\frac{z}{\mu_2} . \quad (3.100)$$

Hieraus folgt die Lösung

$$p(z) = c e^{-\frac{z^2}{2\mu_2}} .$$

Die Integrationskonstante c wird aus der Forderung

$$\int_{-\infty}^{\infty} p(z) dz = 1$$

berechnet und ist

$$c = \frac{1}{\sqrt{2\pi \mu_2}} .$$

Damit erhält man schließlich die bekannte Beziehung für die Normalverteilung mit Ursprung im Mittelwert und Streuung

$$\sigma = \sqrt{\mu_2}$$

$$p(z) = \frac{1}{\sqrt{2\pi \sigma}} e^{-\frac{z^2}{2\sigma^2}} . \quad (3.101)$$

Abschließend wird im Bild 3.12 eine Übersicht über die verschiedenen Kurventypen des Pearsonschen Kurvensystems und über ihre Lage im (β_1, β_2) -Diagramm gegeben.

Auf der Achse $\beta_1 = 0$ sind die Kurven symmetrisch. Bei $\beta_1 = 0$, $\beta_2 = 3$ liegt die Normalverteilung. Darüber sind die Kurven "spitzer", darunter "flacher" als die Normalverteilung. Die Kurven mit (β_1, β_2) -Werten auf der Geraden $\beta_1 = 2$ sind alle für $\mu_3 > 0$ gezeichnet worden. Für $\mu_3 < 0$ wären sie symmetrisch zu diesen.

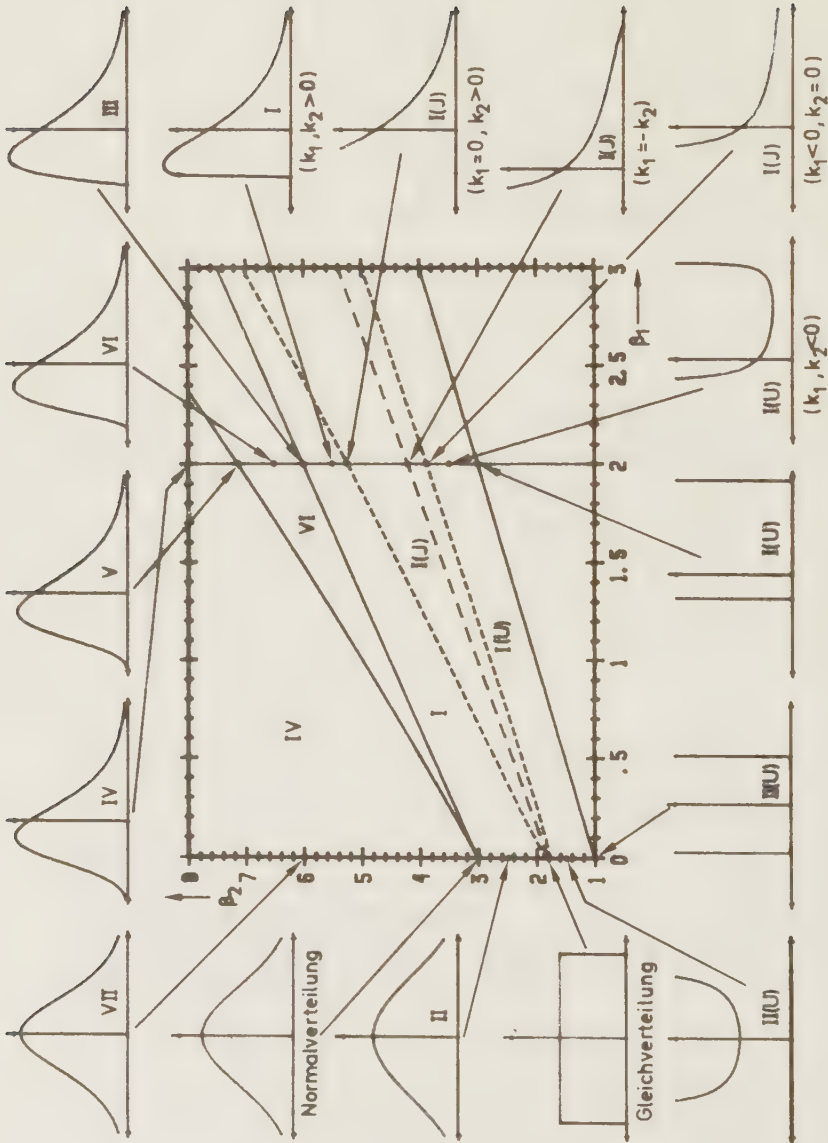


Bild 3.12: Die Pearson'schen Kurven und ihre Lage im (β_1, β_2) -Diagramm

4.3 Berechnung der statistischen Toleranzgrenzen

In den vorangegangenen Abschnitten wurde gezeigt, wie ausgehend von den Momenten bis zur vierten Ordnung, die Wahrscheinlichkeitsdichte einer Zufallsvariable durch Pearsonsche Kurven approximiert werden kann. In dem vorliegenden Fall, in dem die Zufallsvariable eine Netzwerkfunktion ist, interessiert jedoch nicht so sehr die Wahrscheinlichkeitsdichte selbst, sondern vielmehr bestimmte Quantile^{*)} dieser Funktion, d.h. jene Werte der Netzwerkfunktion, innerhalb derer ein gewisser Anteil aller Realisierungen liegt. Diese Quantile werden statistische Toleranzgrenzen (im folgenden einfach Toleranzgrenzen) genannt. Sie stehen immer im Zusammenhang mit einer Zahl Q , die die Wahrscheinlichkeit in % dafür angibt, daß die Realisierungen innerhalb dieser Grenzen liegen (siehe Bild 3.1). Sind z.B. die 90 %-Toleranzgrenzen einer Netzwerkfunktion anzugeben, so sind die Quantile für $q = 0,05$ und $q = 0,95$ zu berechnen.

Gemäß Gl. (A.16) wird ein Quantil z_q der Wahrscheinlichkeitsdichte $p(z)$ durch Auflösen der Integralgleichung

$$\int_{-\infty}^{z_q} p(z) dz = q \quad (3.102)$$

nach z_q berechnet. Soll nun die untere Toleranzgrenze z_u und die obere Toleranzgrenze z_o der Netzwerkfunktion z berechnet werden, so daß Q % aller Realisierungen zwischen diesen Grenzen liegen, so ist in Gl. (3.102) an Stelle von q und z_q

$$q_u = \frac{1}{2} \left(1 - \frac{Q}{100} \right) \quad \text{und} \quad z_u \quad (3.103)$$

bzw.

$$q_o = \frac{1}{2} \left(1 + \frac{Q}{100} \right) \quad \text{und} \quad z_o \quad (3.104)$$

einzusetzen.

^{*)} vgl. Anhang Abschnitt 1.9

Die Schwierigkeit in der Ermittlung der Toleranzgrenzen von Netzwerkfunktionen, deren Verteilung durch eine Kurve des Pearsonschen Systems approximiert wurde, besteht darin, daß die Gl. (3.102) mit den Wahrscheinlichkeitsdichtefunktionen $p(z)$ aus Abschnitt 4.2 nicht geschlossen gelöst werden kann. Gl. (3.102) kann daher nur mit Methoden der numerischen Integration gelöst werden.

4.3.1 Algorithmus für die Bestimmung der Toleranzgrenzen

Bei der numerischen Integration der Gl. (3.102) treten Schwierigkeiten wegen der unendlichen unteren Integrationsgrenze auf. In den Pearsonschen Kurven ist jedoch die Wahrscheinlichkeitsdichte $p(z)$ entweder beidseitig (Typ I und II) oder zumindest einseitig (Typ III, V und VI) beschränkt oder es besteht die Möglichkeit, durch eine geeignete Variablentransformation die unendliche Integrationsgrenze in eine endliche überzuführen. An Stelle von Gl. (3.102) wird daher eine Integralgleichung der Art

$$\int_a^{z_q} p(z) dz = q \quad (3.105)$$

zu lösen sein.

Setzt man die Integrationsgrenze z_q als bekannt voraus, so könnte der Wert q des Integrals auf sehr einfache Weise nach der Simpson-Methode näherungsweise berechnet werden. Hierzu wird das gesamte Integrationsintervall $[a, z_q]$ in eine gerade Zahl n von Intervallen der Breite

$$h = \frac{z_q - a}{n}$$

unterteilt. Es entstehen dadurch $n+1$ Stützstellen $z_0, z_1, \dots, z_n$, wobei $z_0 = a$ und $z_n = z_q$ ist. Damit ist

$$\begin{aligned} q = \int_a^{z_q} p(z) dz &\approx \frac{h}{3} \{ p(z_0) + 4p(z_1) + 2p(z_2) + 4p(z_3) + \dots \\ &+ 2p(z_{n-2}) + 4p(z_{n-1}) + p(z_n) \} \end{aligned} \quad (3.106)$$

Da aber hier nicht q , sondern die obere Integrationsgrenze z_q die Unbekannte ist, kann die Anzahl der Stützstellen und deren Abstand nicht von Anfang an angegeben werden. Zur Lösung der Gl. (3.106) nach z_q wird daher folgendermaßen vorgegangen:

- (i) Es wird ein Stützstellenabstand h angenommen, der klein genug ist, um eine hinreichend genaue Auswertung der numerischen Integration in Gl. (3.106) zu gewährleisten.
- (ii) Der Ausdruck in den großen Klammern in Gl. (3.106) wird nun schrittweise wie folgt berechnet:

$$\begin{aligned} F_1 &= p(z_0) + 4p(z_1) + p(z_2) \\ F_2 &= F_1 + p(z_2) + 4p(z_3) + p(z_4) \\ &\vdots \\ F_i &= F_{i-1} + p(z_{2i-2}) + 4p(z_{2i-1}) + p(z_{2i}). \end{aligned}$$

- (iii) Nach jedem Schritt wird F_i mit $R = 3q/h$ verglichen.
- (iv) Ist $F_i < R$, so bedeutet dies, daß der Wert des Integrals bis zur Stützstelle z_{2i} kleiner ist als q . z_{2i} liegt daher noch unterhalb des gesuchten Quantils z_q . In diesem Fall wird im nächsten Schritt F_{i+1} berechnet (iii).
- (v) Ist jedoch $F_i > R$, so liegt die Stützstelle z_{2i} bereits oberhalb von z_q . In diesem Fall kann der Zyklus abgebrochen werden und das gesuchte Quantil näherungsweise aus der linearen Interpolation zwischen z_{2i} und z_{2i-2} berechnet werden

$$z_q \approx z_{2i-2} + 2h \frac{R - F_{i-1}}{F_i - F_{i-1}}. \quad (3.107)$$

Dieser Algorithmus läßt sich sehr einfach programmieren und eignet sich daher besonders gut für die automatische Bestimmung der Toleranzgrenzen auf dem Digitalrechner.

5. Anwendungsbeispiele

Die Berechnung statistischer Toleranzgrenzen nach der in den vorangegangenen Abschnitten beschriebenen Methode wurde an einer Reihe von praktischen Beispielen erfolgreich erprobt. Einige davon sollen im folgenden detailliert behandelt werden. Die Ergebnisse werden dabei mit denen, die man auf Grund des Fehlerfortpflanzungsgesetzes und der Worst-Case-Methode erhält, verglichen. Um die Güte der Näherung der Wahrscheinlichkeitsdichte mit Hilfe der Pearsonschen Kurven beurteilen zu können, werden diese Kurven an verschiedenen Frequenzpunkten mit dem Histogramm aus einer Monte-Carlo-Analyse verglichen. Hierzu werden hauptsächlich solche Frequenzpunkte ausgesucht, in denen die nach dem Fehlerfortpflanzungsgesetz berechneten Toleranzgrenzen stark von denen abweichen, die aus der Näherung durch Pearsonsche Kurven ermittelt werden.

Die erforderlichen numerischen Rechnungen wurden alle auf einem Rechner vom Typ Hewlett-Packard 21 MX des Lehrstuhls für Allgemeine und Theoretische Elektrotechnik durchgeführt. Die Ergebnisse werden in der Form von Kurven und Histogrammen dargestellt. Diese wurden mit Hilfe eines an den Rechner angeschlossenen Plotters automatisch erstellt.

5.1 1. Beispiel

Als erstes und einfachstes Beispiel wird das in [42] analysierte Netzwerk aus Bild 3.13 herangezogen. Die Induktivität und die beiden Kapazitäten besitzen die im Bild angegebenen Sollwerte, sie haben jeweils eine Toleranz von $\pm 5\%$ und sind gleichverteilt. Die beiden Abschlußwiderstände werden auf ihren Sollwert eingestellt (Toleranz $\pm 0\%$).

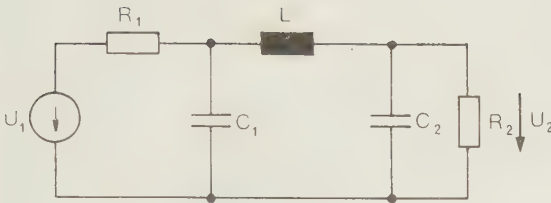


Bild 3.13: Netzwerk des ersten Beispiels

$$L = 0,9487 \text{ H} ; \quad C_1 = C_2 = 2,206 \text{ F} ; \quad R_1 = R_2 = 1 \Omega$$

Die Aufgabe besteht darin, die 98%-Toleranzgrenzen der Dämpfung in Neper

$$a(\omega) = -\ln \left| \frac{u_2}{u_1} \right|$$

zu ermitteln.

Das Ergebnis wird im Bild 3.14 gezeigt. Zum Vergleich sind hier auch die 98%-Toleranzgrenzen nach dem Fehlerfortpflanzungsgesetz und die Worst-Case-Grenzen eingezeichnet worden. Man erkennt, daß diese beiden letzten Verfahren in der Umgebung der Anpassungsstelle bei $\omega = 0,866 \text{ s}^{-1}$ sehr große Fehler liefern. Das Toleranzschema für die Umgebung dieser Anpassungsstelle wurde ausgeblendet und im Bild 3.15 vergrößert wiedergegeben.

Da die Empfindlichkeiten erster Ordnung der Dämpfung in den Anpassungsstellen von Reaktanzfiltern verschwinden, ist die nach dem Fehlerfortpflanzungsgesetz Gl. (3.11) berechnete Varianz gleich Null. Damit verschwinden auch die Toleranzgrenzen von Δa . Aus dem selben Grund sind auch die Worst-Case-Grenzen nach Gl. (3.4) gleich Null. Diese Tatsache ist in den Bildern 3.14 und 3.15 sehr gut zu erkennen.

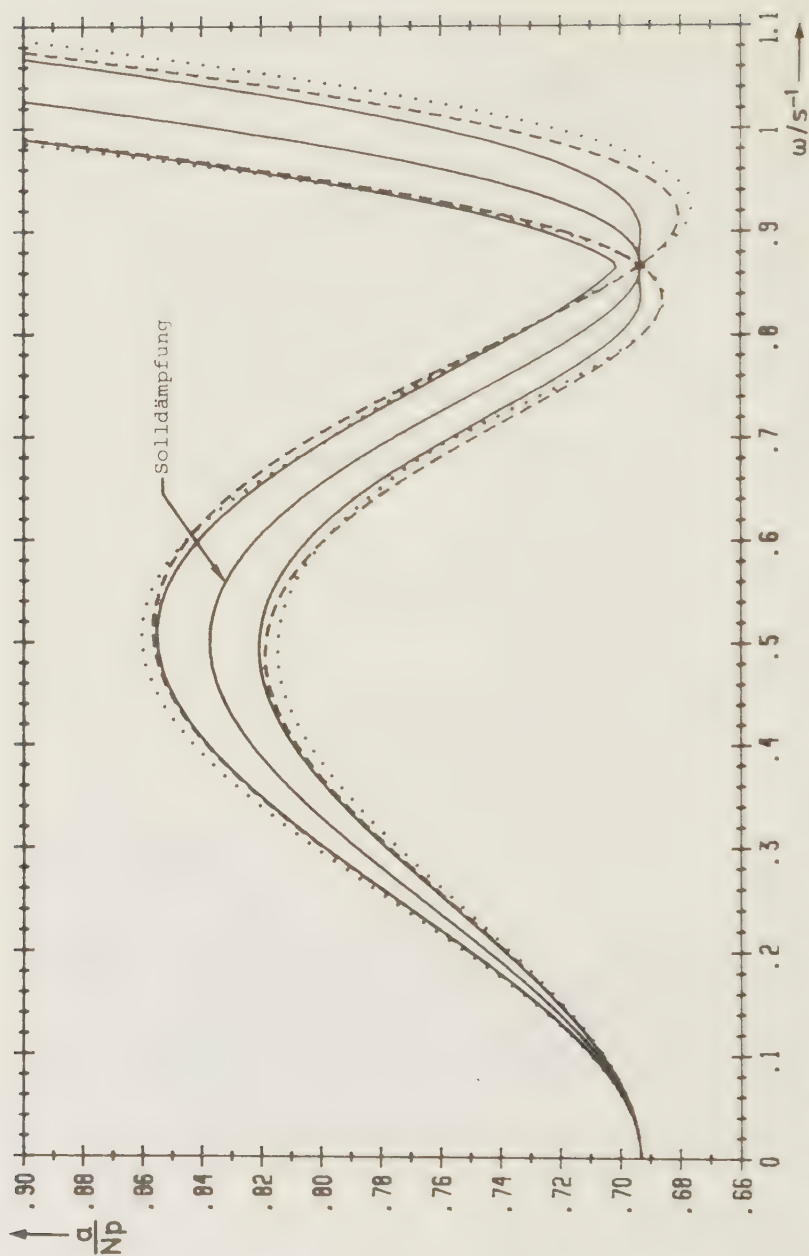


Bild 3.14: 98%-Toleranzgrenzen des Netzwerks aus Bild 3.13. —: Näherung durch Pearson-Kurven; ---: Fehlerfortpflanzungsgesetz; ...: Worst-Case-Grenzen

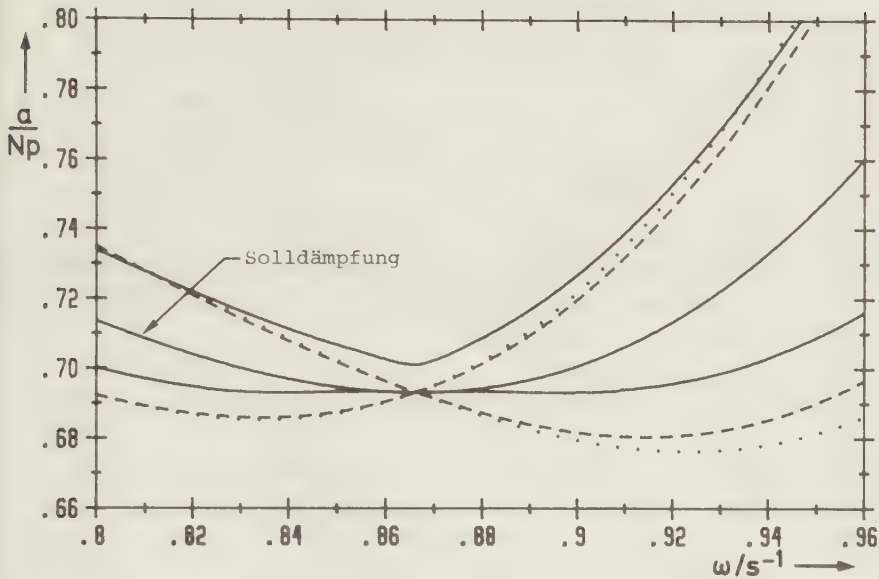


Bild 3.15: Toleranzschema in der Nähe der Anpassungsstelle
 $\omega = 0,866 \text{ s}^{-1}$

- 98 % Toleranzgrenzen mit Näherung durch Pearson-Kurven
- 98 % Toleranzgrenzen nach Fehlerfortpflanzungsgesetz
- Worst-Case-Grenzen

Die mit Hilfe des Pearsonschen Kurvensystems ermittelten Toleranzgrenzen im Bereich der Anpassungsstelle zeigen ein differenziertes Verhalten der oberen und unteren Grenze auf. Während nach oben eine von Null verschiedene Toleranz existiert, stimmt die untere Toleranzgrenze mit der Sollkurve überein. Dies rührt daher, daß an Anpassungsstellen von Reaktanzfiltern, eine Änderung der Bauelementewerte nur eine Erhöhung der Dämpfung hervorrufen kann [32].

Das Bild 3.16 zeigt einen Ausschnitt aus dem (β_1, β_2) -Diagramm in das der Verlauf der (β_1, β_2) -Wertepaare beim Durchschreiten des Frequenzbereiches $0,5 \leq \omega \leq 1,1$ eingezeichnet wurde. Dieses

Bild zeigt auf sehr anschauliche Weise, welche Form die Verteilung in den verschiedenen Frequenzpunkten hat. Zunächst kann man daraus entnehmen, daß die Kurven alle vom Typ I oder II sind. Im Bereich $0,5 \leq \omega \leq 0,7$ sind die Verteilungen annähernd symmetrisch und werden für aufsteigende Frequenzwerte in diesem Bereich zunehmend flacher (da β_2 abnimmt). Dies ist auch in den Bildern 3.17a,b,c sehr gut zu erkennen. Im Bereich $0,7 < \omega < 1,1$ sind die Verteilungen schief ($\beta_1 > 0$) und im Bereich $0,82 < \omega < 0,91$ sind sie außerdem J-förmig. Diese auf Grund des (β_1, β_2) -Diagramms von Bild 3.16 gemachten Aussagen stimmen mit den in Bild 3.18 gezeigten Ergebnissen überein.

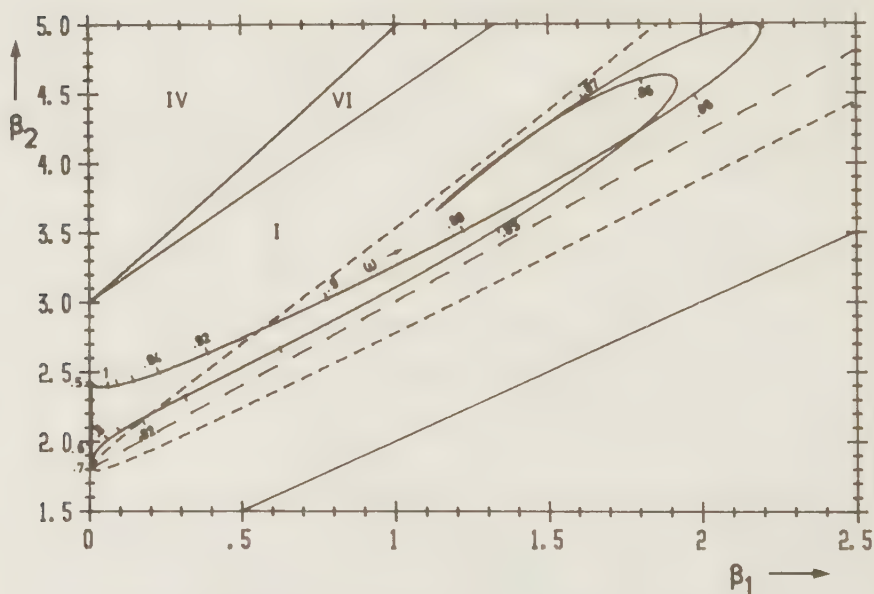


Bild 3.16: Verlauf der (β_1, β_2) -Wertepaare im Frequenzbereich $0,5 \leq \omega \leq 1,1$

In den Bildern 3.17 und 3.18 werden für einige typischen Frequenzpunkte Vergleiche der Wahrscheinlichkeitsdichte $p(\Delta a)$ mit der nach dem Fehlerfortpflanzungsgesetz gewonnenen Normalverteilung und mit dem Histogramm aus einer Monte-Carlo-Analyse durchgeführt. Die durchgehend gezeichnete Kurve stellt dabei

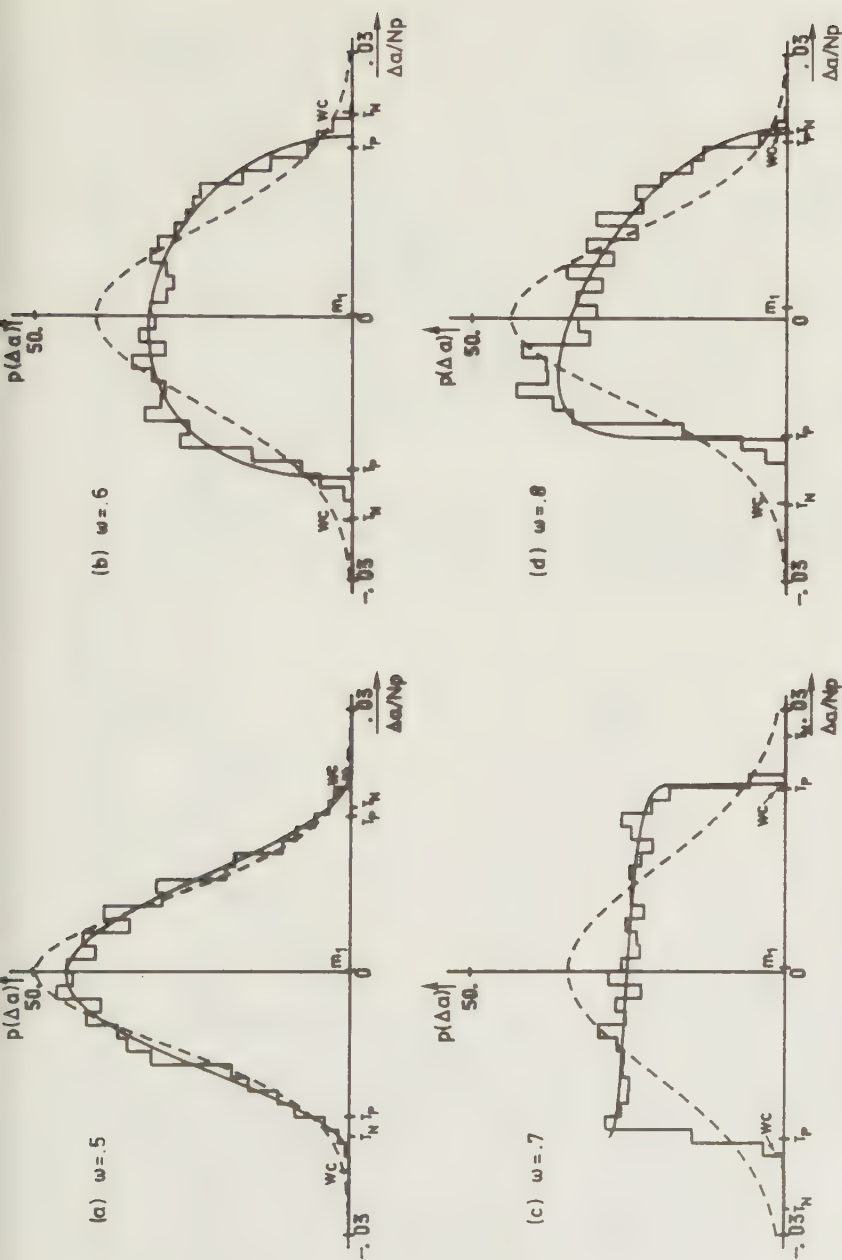


Bild 3.17: Vergleich der Näherung durch Pearson-Kurven mit Fehlerfortpflanzungsgesetz und Monte-Carlo-Analyse (Umfang der Stichprobe: 3000)

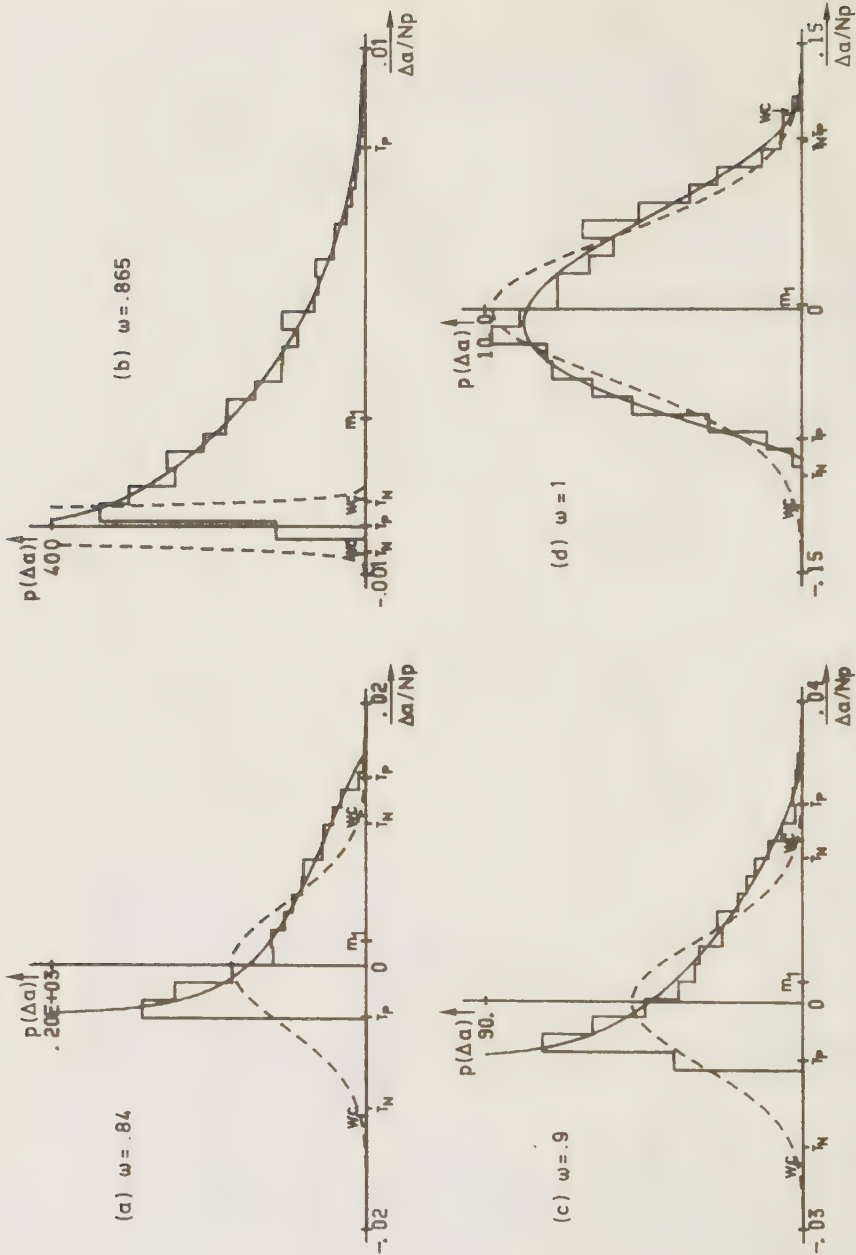


Bild 3.18: Vergleich der Näherung durch Pearson-Kurven mit Fehlerfortpflanzungsgesetz und Monte-Carlo-Analyse (Umfang der Stichprobe: 3000)

die durch eine Pearson-Kurve approximierte Wahrscheinlichkeitsdichte dar und die gestrichelte Kurve die Normalverteilung. Aus diesen Bildern ist zu erkennen, daß die Approximation mit den Pearsonschen Kurven sehr gut mit dem Ergebnis aus der Monte-Carlo-Analyse übereinstimmt. Die Normalverteilung hingegen, weicht in den meisten Fällen sehr stark vom Histogramm der MC-Analyse ab. Dies ist besonders für den Frequenzpunkt $\omega = 0,865$ der Fall (Bild 3.18b). Dieser Frequenzpunkt liegt in unmittelbarer Nähe der Anpassungsstelle bei $\omega = 0,866$. Die Empfindlichkeiten erster Ordnung der Dämpfung gehen hier gegen Null und mit ihnen auch die Varianz der Normalverteilung. Die Form der Verteilung wird daher an dieser Stelle praktisch nur noch von den Empfindlichkeiten zweiter Ordnung bestimmt.

In den Bildern 3.17 und 3.18 sind auf der Abszisse die 98%-Toleranzgrenzen nach beiden Verfahren, die Worst-Case-Grenzen und der Mittelwert der Verteilung eingetragen:

T_P : Toleranzgrenzen aus der Näherung durch Pearsonsche Kurven

T_N : Toleranzgrenzen aus der Normalverteilung nach dem Fehlerfortpflanzungsgesetz

WC: Worst-Case-Grenzen

m_1 : Mittelwert der Pearson-Kurve

5.2 2. Beispiel

Als zweites Beispiel wird das nach Åkerberg und Mossberg [30] realisierte biquadratische Teilglied eines PCM-Tiefpasses mit der Übertragungsfunktion

$$H(p) = \frac{U}{U_1} \cdot \frac{p^2 + 0,907017 \cdot 10^9}{p^2 + 0,522007 \cdot 10^4 p + 0,582528 \cdot 10^9}$$

betrachtet (Bild 3.19).

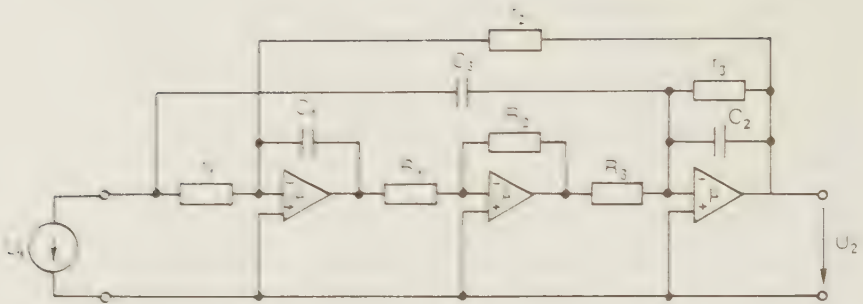


Bild 3.19: Netzwerk des 1. Beispiels

Die Operationsverstärker werden als ideal angenommen ($\mu \rightarrow \infty$) und die passiven Bauelemente haben folgende Sollwerte:

$$r_1 = 3,3204 \text{ k}\Omega, \quad R_1 = 1 \text{ k}\Omega, \quad C_1 = 10 \text{ nF},$$

$$r_2 = 5,1694 \text{ k}\Omega, \quad R_2 = 1 \text{ k}\Omega, \quad C_2 = 10 \text{ nF},$$

$$r_3 = 19,157 \text{ k}\Omega, \quad R_3 = 3,3204 \text{ k}\Omega, \quad C_3 = 10 \text{ nF}.$$

Es sollen die statistischen Toleranzgrenzen der Dämpfung $a(\omega) = -\ln |H(j\omega)|$ berechnet werden, innerhalb derer 97,5 % aller realisierten Schaltungen liegen, wenn alle passiven Netzwerkelemente gleichverteilt sind und eine Toleranz von $\pm 5 \%$ festsetzen.

Das Ergebnis wird im Bild 3.20 gezeigt und mit den Toleranzgrenzen nach dem Fehlerfortpflanzungsgesetz (gestrichelte Kurven) und den Worst-Case-Grenzen (gepunktete Kurven) verglichen. Für Frequenzen oberhalb des Dämpfungsminimums sind erhebliche Abweichungen zwischen den verschiedenen Toleranzkurven festzustellen. Dieser Bereich soll nun näher untersucht werden.

Im Bild 3.21 wird ein Ausschnitt des Toleranzschemas für den Frequenzbereich 1 kHz $\leq f \leq 4$ kHz wiedergegeben und in Bild 3.22 der Ausgangswertverlauf der $(-0,82)$ -Wertepaare gezeigt.

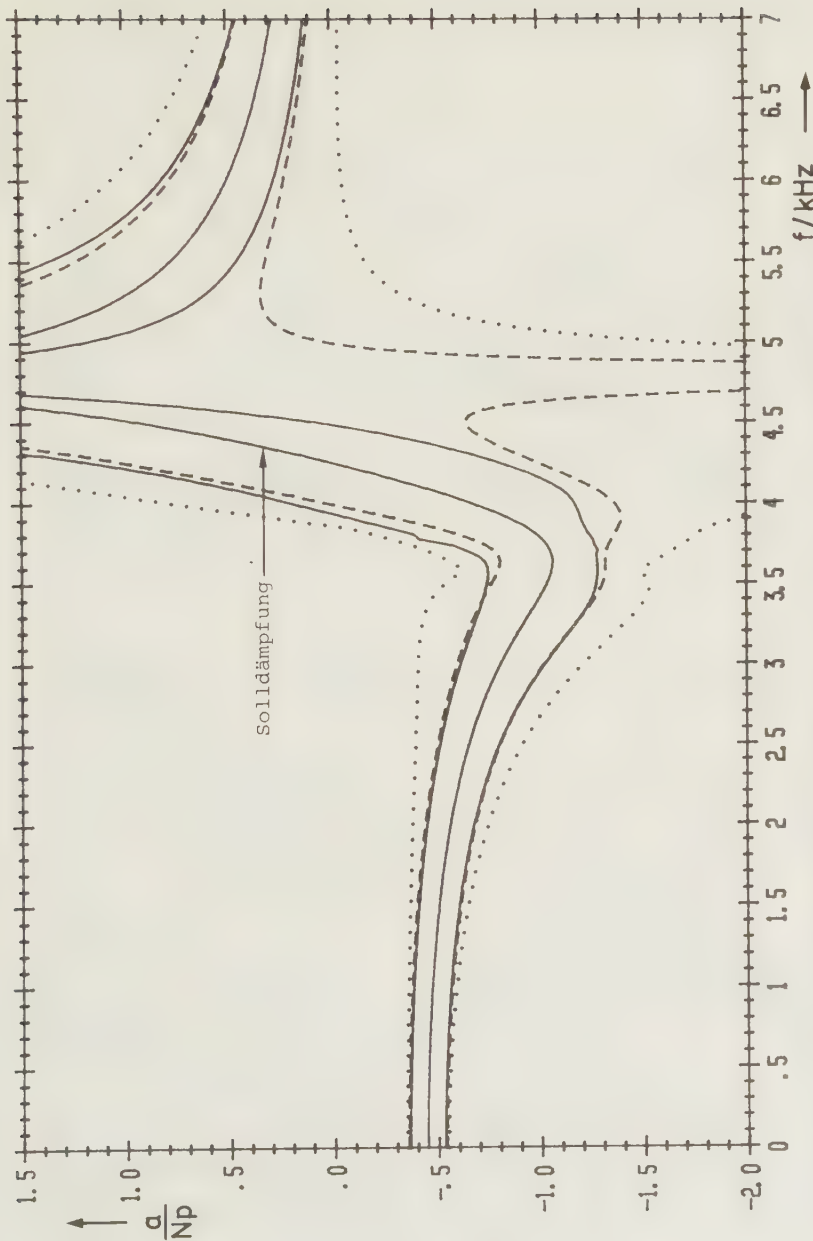


Bild 3.20: 97,5%-Toleranzgrenzen des Netzwerks aus Bild 3.19. —: Näherung durch Pearson-Kurven; ---: Fehlerfortpflanzungsgesetz; ...: Worst-Case-Grenzen

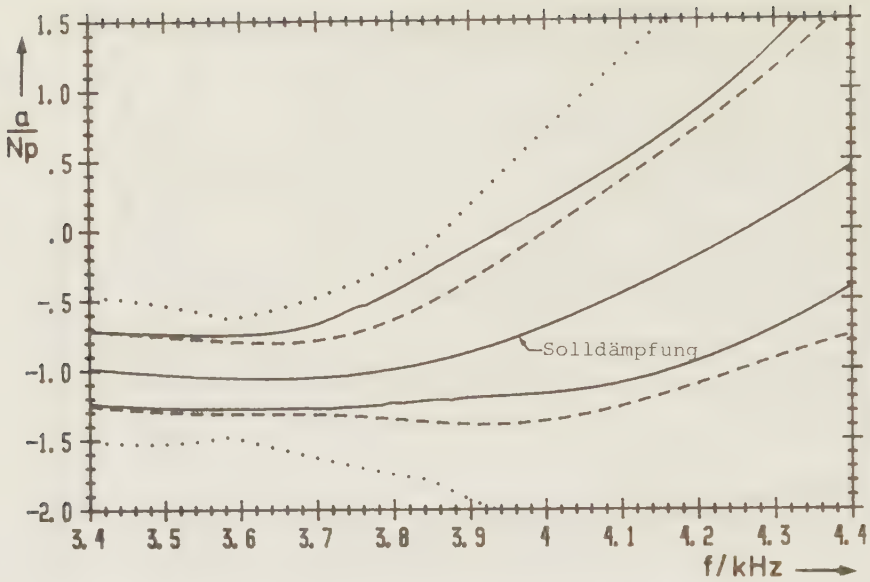


Bild 3.21: Ausschnitt aus dem Toleranzschema aus Bild 3.20

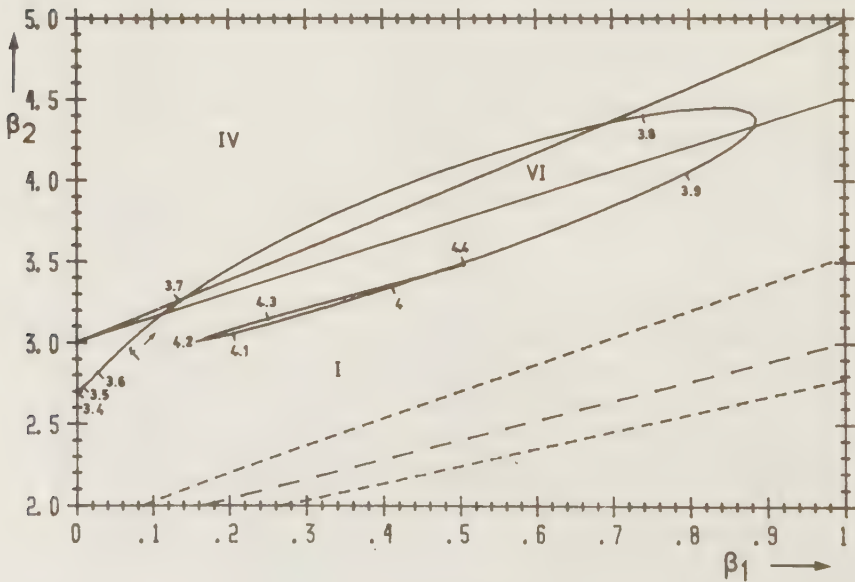


Bild 3.22: Verlauf der (β_1, β_2) -Wertepaare im Frequenzbereich $3,4\text{kHz} \leq f \leq 4,4\text{kHz}$

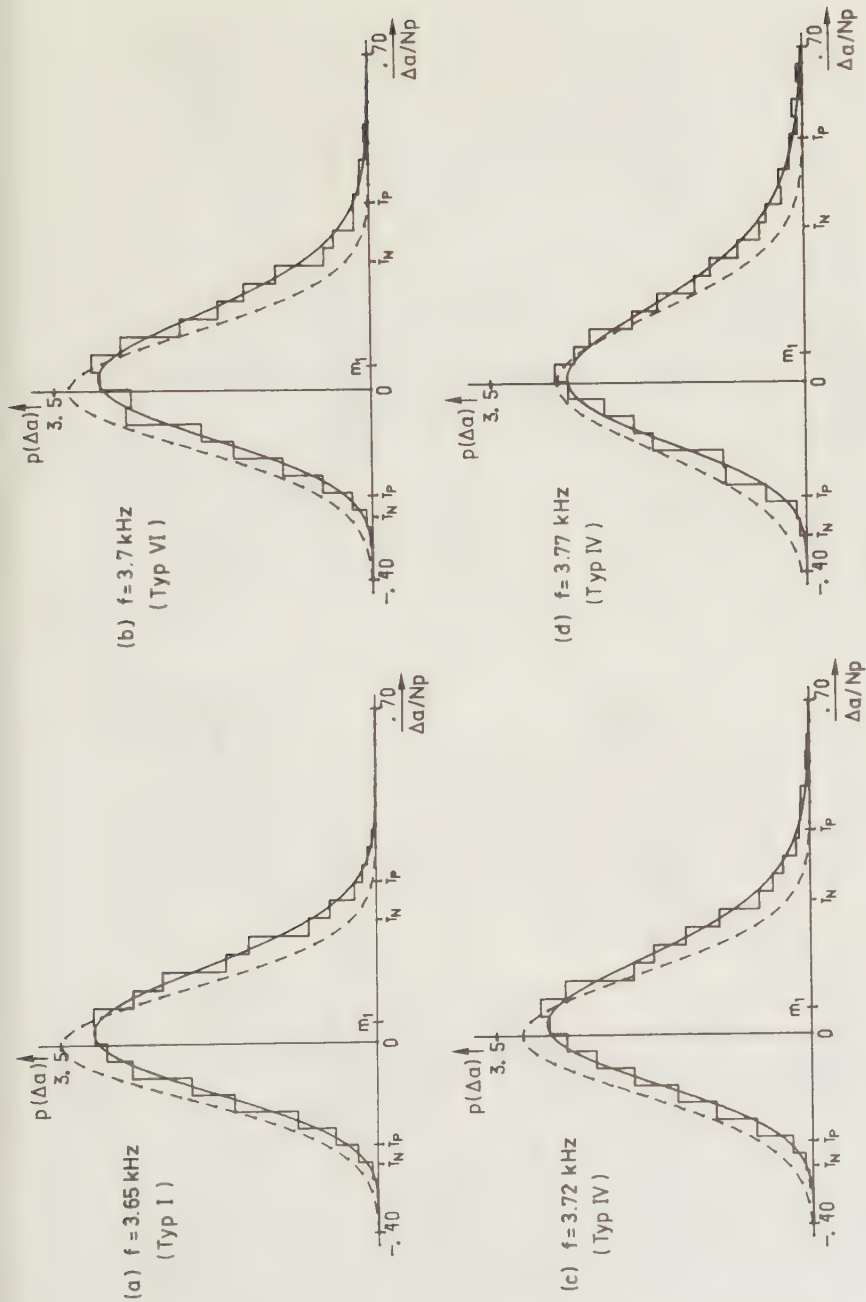


Bild 3.23: Vergleich der Näherung durch Pearson-Kurven mit Fehlerfortpflanzungsgesetz und Monte-Carlo-Analyse (Umfang der Stichprobe: 3000)

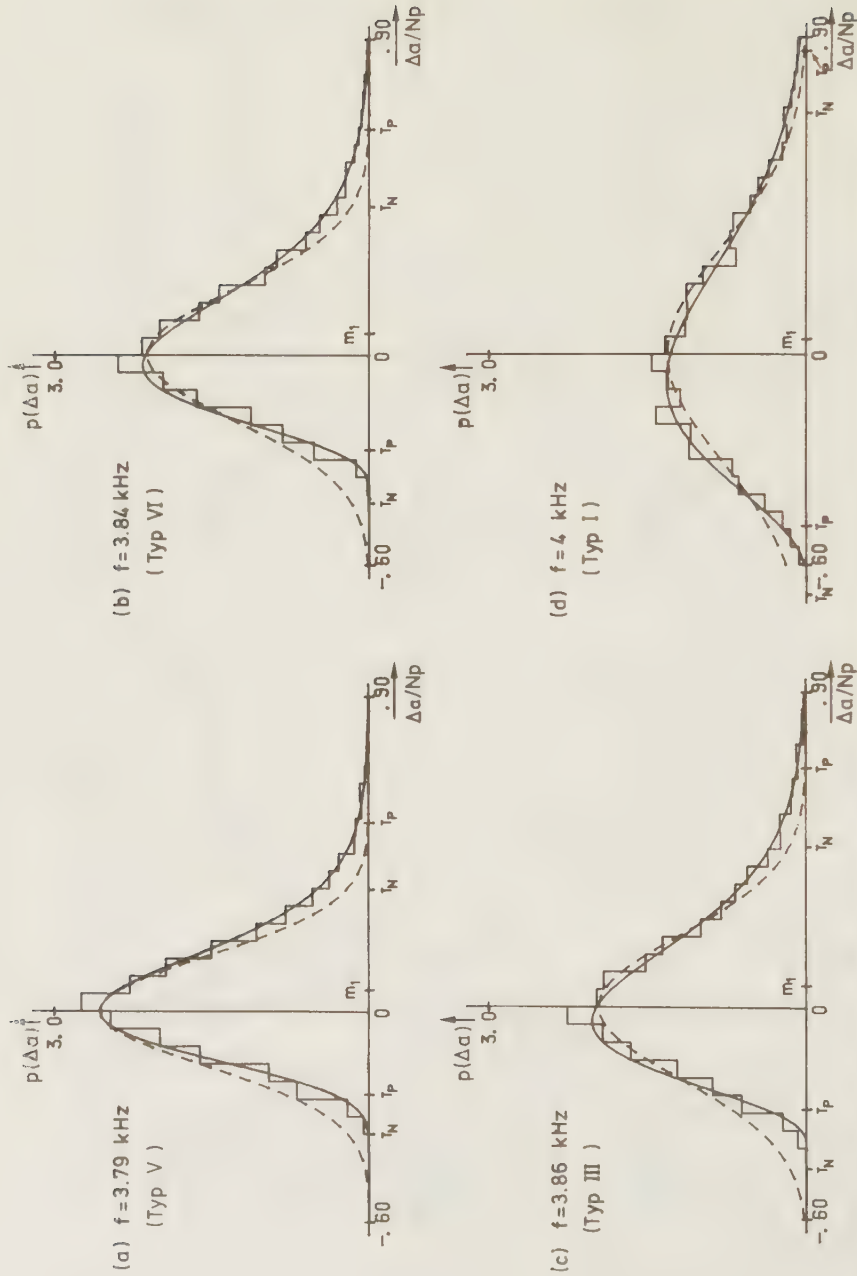


Bild 3.24: Vergleich der Näherung durch Pearson-Kurven mit Fehlerfortpflanzungsgesetz und Monte-Carlo-Analyse (Umfang der Stichprobe: 3000)

Aus diesem letzten Bild ist zu erkennen, daß die Werte von Schiefe und Exzeß auch andere Gebiete als die der Pearson-schen Kurven vom Typ I durchschreiten. Außer den Kurven vom Typ I werden daher in diesem Fall auch die Kurven vom Typ III, IV, V und VI für die Näherung der Wahrscheinlichkeitsdichte benötigt.

In den Bildern 3.23 und 3.24 werden einige Beispiele der Wahrscheinlichkeitsdichte bei verschiedenen Frequenzpunkten gezeigt, bei denen jeweils ein anderer Kurventyp für die Näherung verwendet werden mußte. Wie im ersten Beispiel wird auch hier ein Vergleich mit der Normalverteilung als Ergebnis einer Toleranzberechnung nach dem Fehlerfortpflanzungsgesetz und mit dem Histogramm aus der Monte-Carlo-Analyse durchgeführt. Man erkennt, daß auch in diesem Fall die hier angewandte Methode eine sehr viel bessere Näherung an die Monte-Carlo-Analyse liefert als das Fehlerfortpflanzungsgesetz.

5.3 3. Beispiel

Als drittes und letztes Beispiel soll der Cauer-Parameter-Tiefpaß vom Grad 7 aus Bild 3.25 mit dem Dämpfungsverlauf nach Bild 3.26 betrachtet werden [39]. Die normierten Sollwerte der Netzwerkelemente können aus der Schaltung im Bild 3.25 entnommen werden.

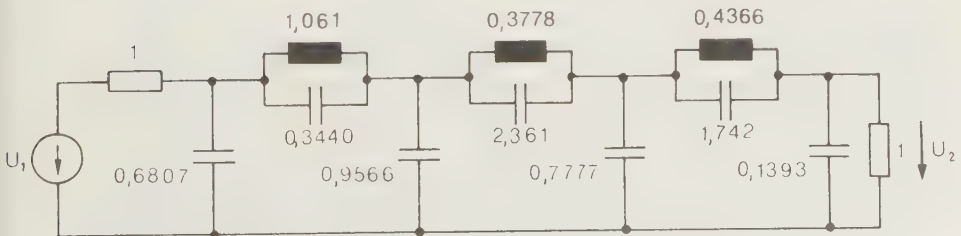


Bild 3.25: Cauer-Parameter-Tiefpaß vom Grad 7

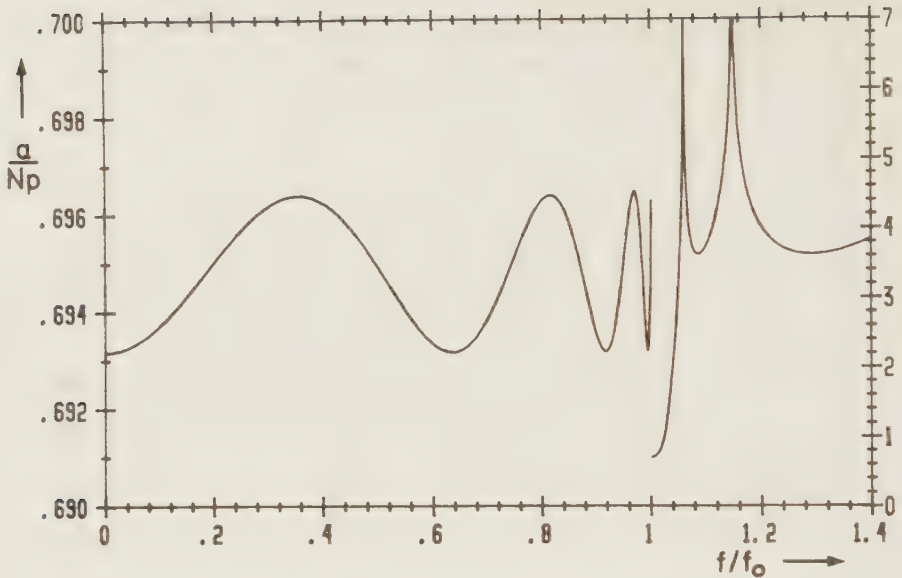


Bild 3.26: Dämpfungsverlauf des Tiefpasses aus Bild 3.25

Es sollen nun die 95 %-Toleranzgrenzen der Dämpfung $a(\omega) = -\ln|U_2/U_1|$ berechnet werden, wenn alle Energie speichernden Elemente gleichverteilt sind und eine Toleranz von $\pm 2 \%$ besitzen. Die beiden Abschlußwiderstände werden auf ihren Sollwert eingestellt (Toleranz $\pm 0 \%$).

Das Ergebnis der Toleranzanalyse wird im Bild 3.27 gezeigt. Hier wird auch ein Vergleich zu den Toleranzgrenzen nach dem Fehlerfortpflanzungsgesetz (gestrichelte Kurven) und den Worst-Case-Grenzen (gepunktete Kurven) gezogen. Man erkennt auch hier wieder starke Abweichungen zwischen den verschiedenen Toleranzkurven in der Umgebung der oberen Anpassungsstellen und in der Nähe der Durchlaßkante. Dieser Bereich wird im Bild 3.28 wiedergegeben. Der dazugehörige Verlauf der (β_1, β_2) -Wertepaare wird im Bild 3.29 gezeigt. Diese Werte bleiben alle im Gebiet der Kurven vom Typ I.

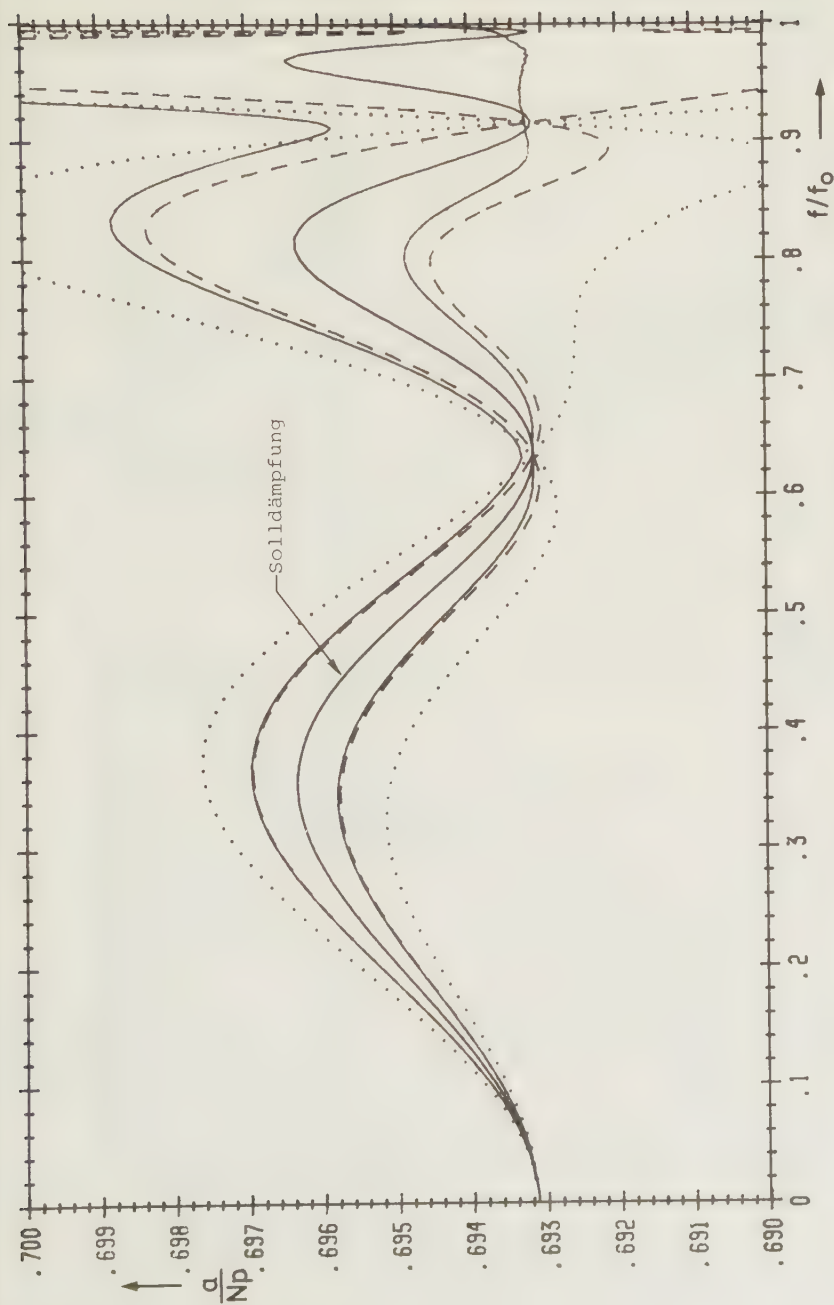


Bild 3.27: 95%-Toleranzgrenzen des Netzwerks aus Bild 3.25. —: Näherung durch Pearson-Kurven; ---: Fehlerfortpflanzungsgesetz; Worst-Case-Grenzen

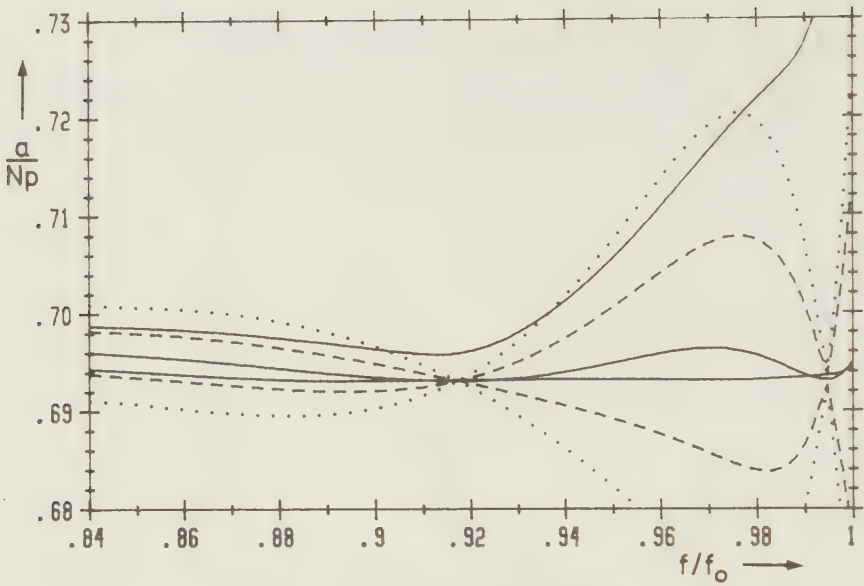


Bild 3.28: Ausschnitt aus dem Toleranzschema aus Bild 3.27

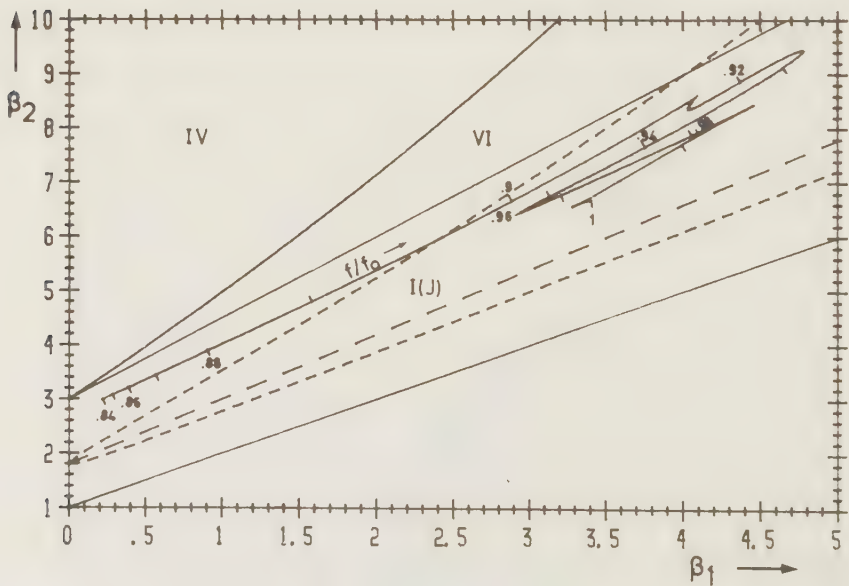


Bild 3.29: Verlauf der (β_1, β_2) -Wertepaare im Frequenzbereich $0,85 \leq f/f_0 \leq 1$

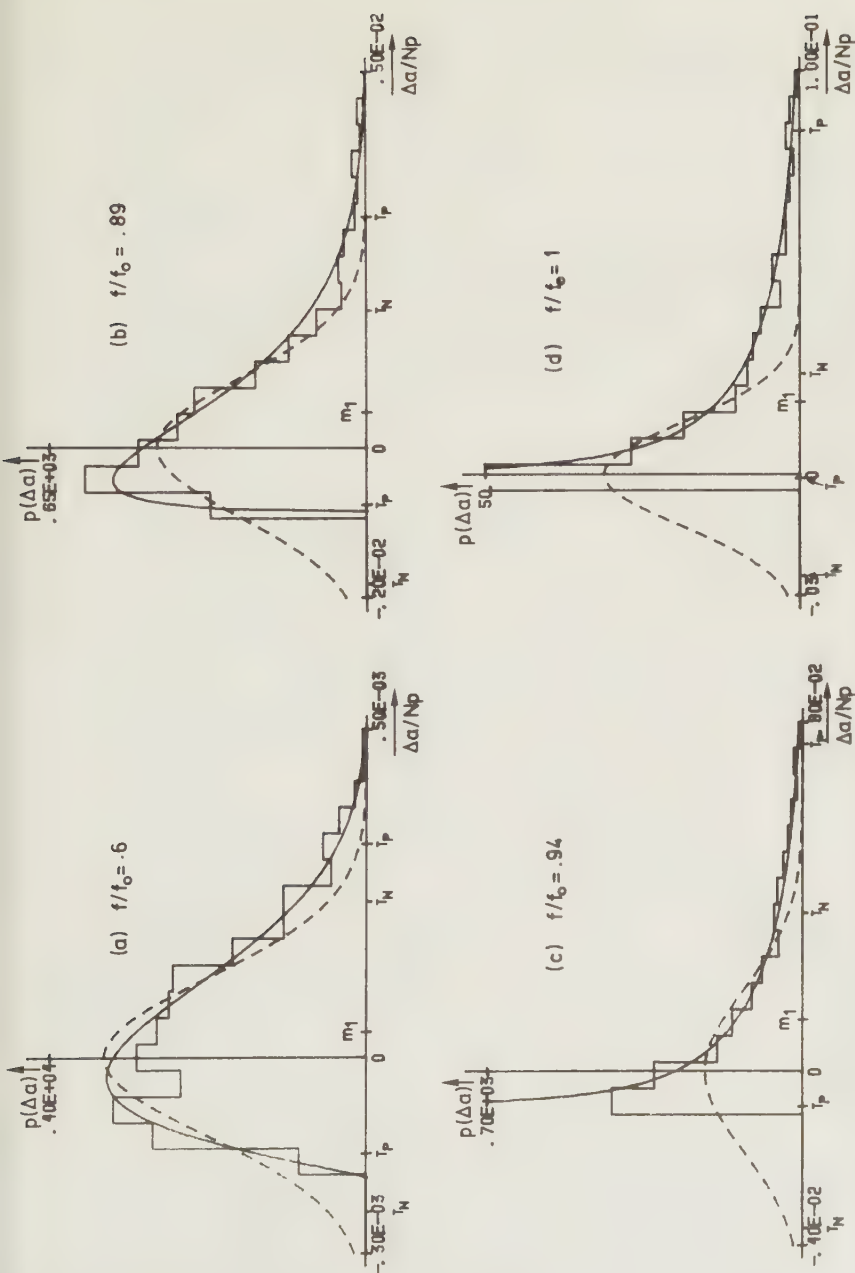


Bild 3.30: Vergleich der Näherung durch Pearson-Kurven mit Fehlerfortpflanzungsgesetz und Monte-Carlo-Analyse (Umfang der Stichprobe: 500)

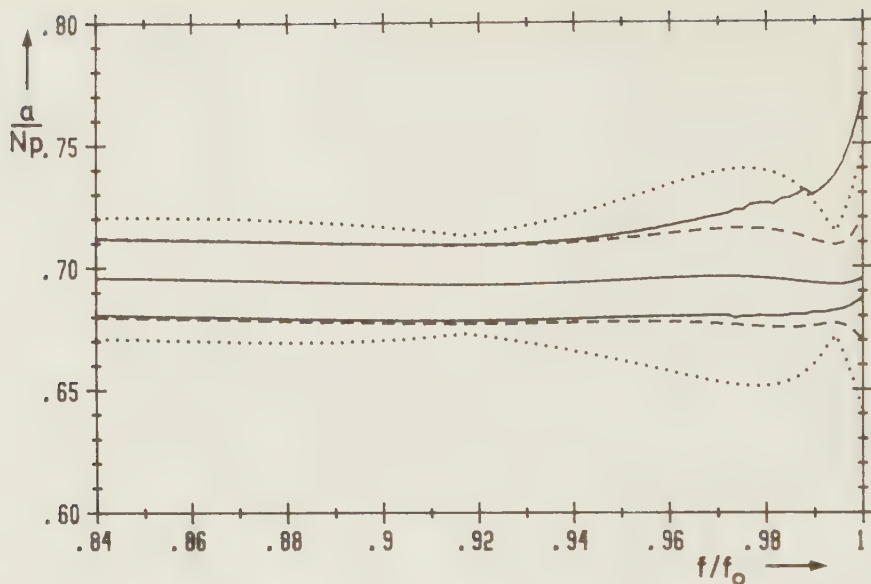


Bild 3.31: 95%-Toleranzgrenzen wenn alle Elemente gleichverteilt sind und eine Toleranz von $\pm 2\%$ besitzen

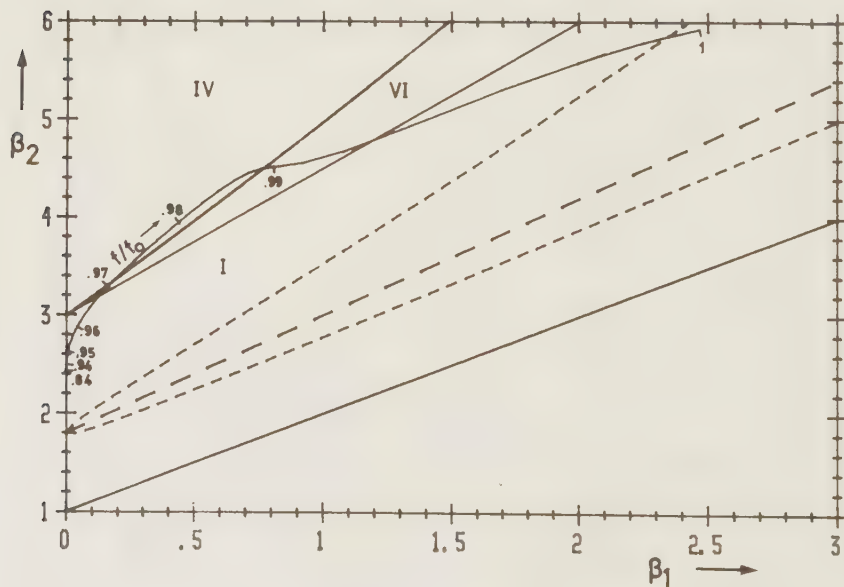


Bild 3.32: Verlauf der (β_1, β_2) -Wertepaare im Frequenzbereich $0,84 \leq f/f_0 \leq 1$

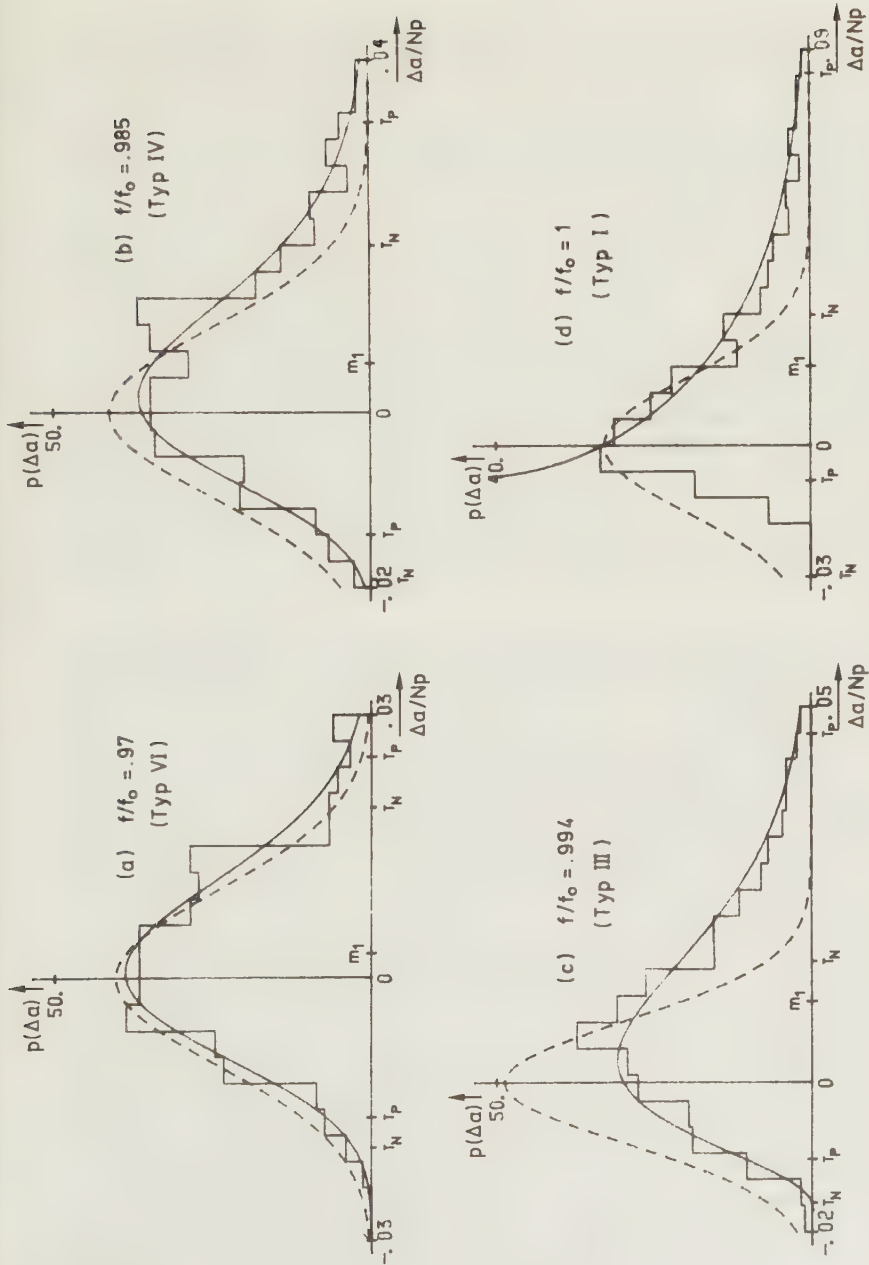


Bild 3.33: Vergleich der Näherung durch Pearson-Kurven mit Fehlerfortpflanzungsgesetz und Monte-Carlo-Analyse (Umfang der Stichprobe: 500)

Um die Güte der Näherung der Wahrscheinlichkeitsdichte durch Pearsonsche Kurven zu zeigen, wurden diese an einigen Frequenzpunkten mit der Monte-Carlo-Analyse verglichen und der Normalverteilung als Ergebnis des Fehlerfortpflanzungsgesetzes gegenübergestellt (Bild 3.30).

Die Toleranzanalyse soll nun auch für den Fall durchgeführt werden, daß sämtliche Netzwerkelemente, also auch die Abschlußwiderstände, eine Toleranz von $\pm 2\%$ (gleichverteilt) besitzen. Das resultierende Toleranzschema für den Frequenzbereich $0,84 \leq f/f_0 \leq 1$ wird im Bild 3.31 und der dazugehörige Verlauf der (β_1, β_2) -Wertepaare im Bild 3.32 gezeigt. Aus dem letzten Bild ist zu entnehmen, daß in diesem Fall auch andere Kurventypen als die vom Typ I für die Näherung der Wahrscheinlichkeitsdichte benötigt werden. Schließlich werden im Bild 3.33 die Verteilungen für einige dieser Frequenzpunkte, in denen unterschiedliche Kurventypen benötigt werden, gezeigt.

Abschließend soll noch einiges über den notwendigen Rechenzeitaufwand für die Toleranzanalyse nach den verschiedenen Verfahren gesagt werden.

Wenn die Toleranzgrenzen aus einer Monte-Carlo-Analyse bestimmt werden, dann sind sie als Zufallsgrößen aufzufassen. Diese Zufallsgrößen sind asymptotisch normalverteilt [45]. Die Standardabweichung dieser Verteilung hängt von dem Umfang der Stichprobe (Anzahl der Versuche in der MC-Analyse) und von der Ausschußrate ab. Je genauer die Bestimmung der Toleranzgrenzen aus der MC-Analyse, d.h. je kleiner die Standardabweichung sein soll, desto größer muß der Umfang der Stichprobe sein. Andererseits muß bei gleichbleibender Genauigkeit der Umfang der Stichprobe um so größer sein, je kleiner die Ausschußrate ist. Da die Standardabweichung der Verteilung der Toleranzgrenze außerdem von dem Wert der Wahrscheinlichkeitsdichte der Netzwerkfunktion für diese Toleranzgrenze abhängt, kann der erforderliche Umfang der Stichprobe bei vorgegebener Ausschußrate und Genauigkeit nicht allgemein angegeben werden. Er ist nämlich nicht nur von Netzwerk zu Netzwerk verschieden, sondern ändert sich im allgemeinen auch innerhalb eines Netzwerkes von einem Frequenzpunkt zum ande-

ren und an einem Frequenzpunkt, von der unteren Toleranzgrenze zur oberen.

Der notwendige Umfang der Stichproben wurde für verschiedene Frequenzpunkte der hier aufgeführten Beispiele berechnet. Unter der Voraussetzung, daß die Toleranzgrenzen mit einer Wahrscheinlichkeit von 95 % auf ± 5 % genau sein sollen, ergab sich dabei im Schnitt ein Stichprobenumfang von ca. 3000.

Vergleichende Zeitmessungen während der Berechnung der drei Beispiele haben gezeigt, daß die Ermittlung der Toleranzgrenzen aus der Monte-Carlo-Analyse mit o.g. Stichprobenumfang im Mittel annähernd um den Faktor 100 langsamer ist als die Berechnung der Toleranzgrenzen aus einer Näherung mit Hilfe des Pearsonschen Kurvensystems. Letzteres Verfahren ist wiederum im Mittel annähernd um den Faktor 10 langsamer als das Verfahren nach dem Fehlerfortpflanzungsgesetz.

Die hier berechneten Beispiele haben gezeigt, daß die Approximation der Wahrscheinlichkeitsdichte durch Pearson-Kurven und die daraus ermittelten Toleranzgrenzen zwar einen höheren Rechenzeitaufwand bedeuten, aber dafür sehr viel genauere Ergebnisse als das Fehlerfortpflanzungsverfahren liefern. Dieser um den Faktor 10 höhere Aufwand an Rechenzeit im Vergleich mit dem Fehlerfortpflanzungsverfahren liegt aber noch in der Nähe von 1/100 dessen, der für die Monte-Carlo-Analyse notwendig wäre.

Z U S A M M E N F A S S U N G

Die vorliegende Arbeit befaßt sich mit der Empfindlichkeitsberechnung elektrischer Netzwerke und ihren Anwendungen, insbesondere der Empfindlichkeitsoptimierung und der Toleranzanalyse.

Im ersten Kapitel werden neben einigen allgemeinen Definitionen der Empfindlichkeiten zwei Verfahren behandelt, mit denen außer den Empfindlichkeiten erster Ordnung auch diejenigen zweiter Ordnung berechnet werden können. Diese Verfahren eignen sich daher besonders gut für die Empfindlichkeitsoptimierung und für die Toleranzanalyse.

Das zweite Kapitel beschäftigt sich mit der Empfindlichkeitsoptimierung. Hier wird zunächst ein geeignetes Gesamtempfindlichkeitsmaß als Zielfunktion für die Optimierung gesucht. Diese Zielfunktion wird durch die einzuhaltenden Nebenbedingungen modifiziert und mit der SUMT-Methode von Fiacco und McCormick minimiert. Die Freiheitsgrade bei der Wahl der Netzwerkparameter können jedoch nicht immer für eine Empfindlichkeitsoptimierung genutzt werden. Aus diesem Grund wird im Kapitel II eingehend die Frage behandelt, wann eine Empfindlichkeitsoptimierung überhaupt möglich ist.

Das letzte Kapitel widmet sich der Toleranzanalyse elektrischer Netzwerke. Hier wird ein Verfahren behandelt, das von der Taylorentwicklung zweiter Ordnung der Netzwerkfunktion ausgeht. Aus dieser Entwicklung und den als bekannt vorausgesetzten Verteilungen der Bauelemente werden die Momente bis zur vierten Ordnung und daraus wiederum Schiefe und Exzeß der Verteilung der Netzwerkfunktion berechnet. Die Werte dieser beiden letzten Parameter bestimmen die für die Näherung zu verwendende Kurve aus dem Pearsonschen Kurvensystem. Aus der somit gewonnenen Näherung der Wahrscheinlichkeitsdichte lassen sich schließlich mit Hilfe numerischer Integrationsverfahren die statistischen Toleranzgrenzen der Netzwerkfunktion berechnen. Das Verfahren läßt sich immer dann anwenden, wenn die Netzwerkfunktion hinreichend genau durch eine Näherung zweiter Ordnung dargestellt

werden kann und wenn für die Beschreibung der Verteilung die Momente bis zur vierten Ordnung ausreichend sind.

Abschließend wird in einigen praktischen Beispielen dieses Verfahren mit dem Worst-Case-, dem Fehlerfortpflanzungs- und dem Monte-Carlo-Verfahren verglichen. Dabei zeigt sich, daß die Näherungen durch das Pearsonsche Kurvensystem sehr gut mit den Ergebnissen aus der Monte-Carlo-Analyse übereinstimmen. Die aus den Pearson-Kurven berechneten Toleranzgrenzen liefern daher viel genauere Ergebnisse als das Fehlerfortpflanzungsgesetz. Obwohl damit der notwendige Aufwand an Rechenzeit gestiegen ist, liegt er noch erheblich unter dem, der für eine Monte-Carlo-Analyse erforderlich wäre.

A N H A N G

Grundbegriffe und Beziehungen der Wahrscheinlichkeitsrechnung |37|, |44|, |45|

Hier werden die im Rahmen dieser Arbeit benötigten Begriffe und Beziehungen der Wahrscheinlichkeitsrechnung in übersichtlicher Form zusammengefaßt. Zufallsgrößen werden dabei im allgemeinen durch große Buchstaben bezeichnet. Für die möglichen Werte dieser Zufallsgrößen werden die entsprechenden kleinen Buchstaben verwendet.

1. Eindimensionale Zufallsvariablen

1.1 Die Wahrscheinlichkeitsdichtefunktion

(häufig auch Wahrscheinlichkeitsdichte, Dichtefunktion, oder einfach Verteilung genannt) $p(x)$ einer Zufallsgröße X ist eine positive Funktion, deren Integral über ein bestimmtes Intervall die Wahrscheinlichkeit dafür ist, daß X in dieses Intervall fällt

$$W(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} p(x) dx . \quad (A.1)$$

Aus dieser Beziehung ist sofort zu erkennen, daß $p(x)$ die Eigenschaft

$$\int_{-\infty}^{\infty} p(x) dx = 1 \quad (A.2)$$

haben muß.

1.2 Die Wahrscheinlichkeitsfunktion (oder einfach Verteilungsfunktion)

$P(x)$ der Zufallsgröße X ist die Wahrscheinlichkeit dafür, daß

bei einer Realisierung von X deren Wert $\leq x$ ist. Sie ergibt sich aus $p(x)$ durch Integration

$$P(x) = W(X \leq x) = \int_{-\infty}^x p(x) dx . \quad (A.3)$$

Die Wahrscheinlichkeit, daß der Wert von X zwischen x_1 und x_2 fällt, ist demnach (vgl. Gl. (A.1))

$$W(x_1 \leq X \leq x_2) = P(x_2) - P(x_1) = \int_{x_1}^{x_2} p(x) dx . \quad (A.4)$$

1.3 Erwartungswert

oder wahrscheinlichkeitstheoretischer Mittelwert der zufälligen Veränderlichen $g(X)$ *) ist die Mittelung über alle möglichen Werte $g(x)$, wobei diese das Gewicht $p(x)$ tragen

$$E(g(X)) = \overline{g(x)} = \int_{-\infty}^{\infty} g(x) p(x) dx . \quad (A.5)$$

Hieraus lassen sich sofort folgende Eigenschaften des Erwartungswertes ableiten:

$$E(k) = k \quad (A.6)$$

$$E\left(\sum_{i=1}^n k_i g_i(X)\right) = \sum_{i=1}^n k_i E(g_i(X)) , \quad (A.7)$$

wobei k und k_i ($i = 1, 2, \dots, n$) Konstanten sind.

1.4 Momente einer Verteilung (Momente um den Nullpunkt)

Der Mittelwert der Funktion $g(X) = X^k$ wird mit Moment k -ter Ordnung der Zufallsvariablen X bezeichnet. Gemäß Gl. (A.5) ist dann

*) Die Funktion $g(X)$ der Zufallsvariablen X ist selbst eine Zufallsvariable.

$$m_k = E(X^k) = \overline{x^k} = \int_{-\infty}^{\infty} x^k p(x) dx . \quad (A.8)$$

Für $k = 1$ erhält man das Moment erster Ordnung oder Mittelwert der Zufallsvariablen X

$$m_1 = E(X) = \bar{x} = \int_{-\infty}^{\infty} x p(x) dx . \quad (A.9)$$

1.5 Die zentralen Momente

Werden die Momente auf den Mittelwert m_1 bezogen, so erhält man die zentralen Momente

$$\mu_k = E(X - E(X))^k = \int_{-\infty}^{\infty} (x - m_1)^k p(x) dx . \quad (A.10)$$

Aus dieser Gleichung können für $k = 1, 2, 3, 4$ folgende nützlichen Beziehungen zwischen den zentralen Momenten μ_k und den Momenten um den Ursprung aufgestellt werden:

$$\mu_1 = 0 , \quad (A.11a)$$

$$\begin{aligned} \mu_2 &= E(X^2) - \{E(X)\}^2 \\ &= m_2 - m_1^2 , \end{aligned} \quad (A.11b)$$

$$\begin{aligned} \mu_3 &= E(X^3) - 3E(X)E(X^2) + 2\{E(X)\}^3 \\ &= m_3 - 3m_1m_2 + 2m_1^3 , \end{aligned} \quad (A.11c)$$

$$\begin{aligned} \mu_4 &= E(X^4) - 4E(X)E(X^3) + 6\{E(X)\}^2E(X^2) - 3\{E(X)\}^4 \\ &= m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4 . \end{aligned} \quad (A.11d)$$

Aus Gl. (A.10) ist außerdem sofort zu erkennen, daß für symmetrische Verteilungen die zentralen Momente ungerader Ordnung verschwinden.

1.6 Dispersion oder Varianz und Streuung

Das zentrale Moment zweiter Ordnung oder die mittlere quadratische Abweichung der zufälligen Veränderlichen X von ihrem Mittelwert m_1 wird Dispersion oder Varianz genannt und mit σ^2 bezeichnet

$$\sigma^2 = \mu_2 = E(X^2) - (E(X))^2. \quad (\text{A.12})$$

Der Wert $\sigma = \sqrt{\mu_2}$ wird Streuung oder Standardabweichung genannt.

1.7 Schiefte

Das zentrale Moment dritter Ordnung kann als Maß für die Schiefe bzw. Asymmetrie der Verteilung einer Zufallsvariablen verwendet werden. Häufiger wird jedoch die relative Größe

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} \quad (\text{A.13})$$

dafür verwendet. Für eine symmetrische Verteilung ist $\mu_3 = 0$ und folglich auch $\beta_1 = 0$.

1.8 Exzeß

Das zentrale Moment vierter Ordnung kann als Maß für den Exzeß einer Verteilung verwendet werden. Häufiger wird jedoch die relative Größe

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \quad (\text{A.14})$$

dafür herangezogen. Man kann zeigen, daß für die Gaußsche Normalverteilung $\beta_2 = 3$ gilt. Daher sind Verteilungen mit $\beta_2 > 3$ "spitzer" und solche mit $\beta_2 < 3$ "flacher" als die Normalverteilung.

1.9 Quantile

Eine weitere Art Informationen über die Verteilung einer zufälligen Veränderlichen zusammenzufassen, ist die Angabe von Quantilen. Mit Quantil q -ter Ordnung wird der Wert x_q der Zufallsvariable X bezeichnet für den die Wahrscheinlichkeitsfunktion den Wert q hat.

x_q wird daher aus der Gleichung

$$P(x_q) = q \quad (\text{A.15})$$

oder (vgl. Gl. (A.3)) aus der Gleichung

$$\int_{-\infty}^{x_q} p(x) \, dx = q \quad (\text{A.16})$$

bestimmt.

Die Kenntnis der Quantile für verschiedene geeignet gewählte Werte von q kann einen guten Überblick über Lage und Streuung der Verteilung geben.

2. Mehrdimensionale zufällige Veränderliche

Um das Wesentliche deutlich hervorzuheben, wird hier hauptsächlich der zweidimensionale Fall behandelt. Der mehrdimensionale Fall läßt sich daraus einfach durch Erweiterung ableiten.

2.1 Die Verbund-Wahrscheinlichkeitsdichtefunktion

(oder Verbund-Dichtefunktion)

$p(x,y)$ wird, analog zum eindimensionalen Fall, durch die Beziehung

$$W(x_1 \leq x \leq x_2, y_1 \leq y \leq y_2) = \int_{y_1}^{y_2} \int_{x_1}^{x_2} p(x,y) dx dy \quad (A.17)$$

definiert. Auch hier gelten die Bedingungen

$$p(x,y) \geq 0 \quad \text{für alle } x,y \quad (A.18a)$$

und

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x,y) dx dy = 1. \quad (A.18b)$$

2.2 Die Verbund-Wahrscheinlichkeitsfunktion

(oder Verbund-Verteilungsfunktion)

$P(x,y)$ wird durch die Beziehung

$$P(x,y) = W(X \leq x, Y \leq y) = \int_{-\infty}^y \int_{-\infty}^x p(x,y) dx dy \quad (A.19)$$

definiert.

2.3 Die Randdichtefunktion

einer zweidimensionalen Zufallsvariablen (X,Y) ist die Wahrscheinlichkeitsdichte einer der beiden Zufallsvariablen. Sie berechnet sich aus der Verbund-Dichtefunktion durch Integration über den Wertebereich der anderen Variablen

$$p_1(x) = \int_{-\infty}^{\infty} p(x,y) dy \quad (\text{A.20a})$$

$$p_2(y) = \int_{-\infty}^{\infty} p(x,y) dx . \quad (\text{A.20b})$$

2.4 Die Randverteilungsfunktion

wird ähnlich wie die Randdichtefunktion definiert. So ist

$$P_1(x) = W(X \leq x, Y < \infty) = \int_{-\infty}^x \int_{-\infty}^{\infty} p(x,y) dy dx = \int_{-\infty}^x p_1(x) dx \quad (\text{A.21a})$$

$$\text{und } P_2(y) = W(X < \infty, Y \leq y) = \int_{-\infty}^{\infty} \int_{-\infty}^y p(x,y) dx dy = \int_{-\infty}^y p_2(y) dy \quad (\text{A.21b})$$

2.5 Unabhängige zufällige Veränderliche

Die Zufallsvariablen X und Y sind unabhängig, wenn für beliebige reelle Zahlenpaare (x,y) die Beziehung

$$P(x,y) = P_1(x) P_2(y) \quad (\text{A.22a})$$

oder, was gleichbedeutend ist, .

$$p(x,y) = p_1(x) p_2(y) \quad (\text{A.22b})$$

gilt.

2.6 Mittelwert oder Erwartungswert einer Funktion einer mehrdimensionalen Zufallsvariablen

- (a) Für den zweidimensionalen Fall ist der Mittelwert der Funktion $g(x,y)$

$$E\{g(X,Y)\} = \overline{g(x,y)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) p(x,y) dx dy . \quad (A.23)$$

Für den Sonderfall, daß $g(x,y) = x$ ist, vereinfacht sich Gl. (A.23) mit Hilfe von Gl. (A.20a) zu

$$E(X) = \int_{-\infty}^{\infty} x p_1(x) dx \quad (A.24)$$

Für den Fall, daß $g(x,y) = x+y$ ist, wobei X und Y abhängige oder unabhängige zufällige Veränderliche sind, erhält man aus Gl. (A.23) mit Hilfe der Gln. (A.20a,b) und (A.24)

$$\begin{aligned} E(X+Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y) p(x,y) dx dy \\ &= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} p(x,y) dy dx + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} p(x,y) dx dy \\ &= \int_{-\infty}^{\infty} x p_1(x) dx + \int_{-\infty}^{\infty} y p_2(y) dy \end{aligned}$$

$$E(X+Y) = E(X) + E(Y) . \quad (A.25)$$

Für den Fall, daß $g(x,y) = xy$ ist, wobei aber nun X und Y unabhängige Zufallsvariablen sind, erhält man auf ähnliche Weise aus Gl. (A.23) mit Hilfe der Gln. (A.22b) und (A.24)

$$E(XY) = E(X) E(Y) . \quad (A.26)$$

- (b) Es soll nun der mehrdimensionale Fall betrachtet werden. Ist $z = g(x_1, x_2, \dots, x_n)$ eine Funktion der Zufallsvariablen $X_1, X_2, \dots, X_n$, dann gilt

$$E(Z) = \bar{z} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n . \quad (A.27)$$

Ähnlich wie in Gl. (A.25) kann auch hier gezeigt werden, daß der Erwartungswert einer Summe endlich vieler zufälliger Veränderlichen gleich der Summe der Erwartungswerte dieser zufälligen Veränderlichen ist

$$E\left\{\sum_{i=1}^n k_i X_i\right\} = \sum_{i=1}^n \{k_i E(X_i)\} . \quad (\text{A.28})$$

Desgleichen kann, ähnlich wie für den zweidimensionalen Fall, Gl. (A.26), gezeigt werden, daß der Erwartungswert des Produktes von endlich vielen unabhängigen Zufallsvariablen gleich dem Produkt der Erwartungswerte dieser Zufallsvariablen ist

$$E\left\{\prod_{i=1}^n X_i\right\} = \prod_{i=1}^n \{E(X_i)\} . \quad (\text{A.29})$$

3.7 Momente einer Funktion einer mehrdimensionalen Zufallsvariablen

Das Moment k -ter Ordnung um den Nullpunkt der Funktion

$z = g(x_1, x_2, \dots, x_n)$ ist

$$m_k = E(Z^k) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z^k p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n . \quad (\text{A.30})$$

Für $k = 1$ erhält man den Mittelwert $m_1 = E(Z)$ und die Gl. (A.30) geht über in die Gl. (A.27).

Sind die Zufallsvariablen $X_1, X_2, \dots, X_n$ voneinander unabhängig, so ist auf Grund von Gl. (A.22b)

$$m_k = E(z^k) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z^k p_1(x_1) p_2(x_2) \dots p_n(x_n) dx_1 dx_2 \dots dx_n . \quad (\text{A.31})$$

2.8 Kovarianz und Korrelation

In ähnlicher Weise wie die Varianz einer eindimensionalen Zufallsvariablen kann auch die Kovarianz zwischen zwei Zufallsvariablen X und Y definiert werden:

$$\text{Cov}(X,Y) = E\{ \{X-E(X)\} \{Y-E(Y)\} \} . \quad (\text{A.32})$$

Nach Ausmultiplizieren und Bilden des Erwartungswertes erhält man daraus

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) . \quad (\text{A.33})$$

Der Korrelationskoeffizient wird wie folgt definiert:

$$\rho_{XY} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} . \quad (\text{A.34})$$

Sind die beiden Zufallsvariablen unabhängig, dann verschwindet wegen Gl. (A.26) sowohl die Kovarianz als auch der Korrelationskoeffizient.

LITERATURVERZEICHNIS

- |1| Karafin, B.J.:
The Optimum Assignment of Component Tolerances for
Electrical Networks.
The Bell System Technical Journal 50 (1971), S.1225-1242

- |2| Pinel, J.F.; Roberts, K.A.:
Tolerance Assignment in Linear Networks Using Nonlinear
Programming.
IEEE Trans. CT-19 (1972), S. 475 - 480

- |3| Huber, B.:
Das Empfindlichkeitsverhalten elektrischer Netzwerke
unter dem Gesichtspunkt minimaler Gesamtempfindlichkeit.
Dissertation, Univ. Erlangen-Nürnberg, 1973

- |4| Mitra, S.K.:
Analysis and Synthesis of Linear Active Networks.
John Wiley, New York 1969

- |5| Heinlein, W.E.; Holmes, W.H.:
Active Filters for Integrated Circuits.
R. Oldenbourg, München 1974

- |6| Bode, H.:
Network Analysis and Feedback Amplifier Design.
Van Nostrand, New York, 1945

- |7| Goldstein, A.J.; Kuo, F.F.:
Multiparameter Sensitivity
IRE Trans. CT-8 (1961), S. 177 und 178

- |8| Geher, K.:
Theory of Network Tolerances.
Akadémiai Kiadó, Budapest 1971

- |9| Butler, W.J.; Haykin, S.S.:
Multiparameter Sensitivity problems in network theory.
Proc. IEE 117 (1970), S. 2228-2236

- [10] Director, S.W.; Rohrer, R.A.:
Automated Network Design - The Frequency-Domain Case.
IEEE Trans. CT-16 (1969), S. 330 - 337
- [11] Bykhovski, M.L.:
Sensitivity and Dynamic Accuracy of Control Systems.
In: Sensitivity Methods in Control Theory.
L. Randanovic, Pergamon Press, London 1966
- [12] Kiss, D.:
Determination of Element Sensitivity Without Derivation.
Proc. of the International Symposium on Network Theory,
Belgrad 1968, S. 617 - 627
- [13] Mucha, J.:
Zwei Verfahren zur Empfindlichkeitsberechnung von Netzwerkfunktionen mit Hilfe eines Digitalrechners.
NTZ 22 (1969), S. 637 - 642
- [14] Unbehauen, R.:
Elektrische Netzwerke.
Springer, Berlin 1972
- [15] Wehrhahn, E.:
Netzwerk-Analyse-Programm ERLNAP.
Programmbeschreibung, Institut für Allgemeine und Theoretische Elektrotechnik der Universität Erlangen-Nürnberg
- [16] Temple, V.A.K.; Butler, W.J.:
Program for Obtaining Exact Sensitivity Functions.
Electronic Letters 6 (1970), S. 58 - 60
- [17] Yokomoto, C.F.:
A Simple Bookkeeping Scheme for Computing Sensitivities of Symbolic Transfer Functions.
IEEE Trans. CAS-21 (1974), S. 606 - 608
- [18] Hakimi, S.L.; Cruz, J.B.:
Measures of Sensitivity for Linear Systems with Large

Multiple Parameter Variations.

IRE Wescon Convention Record, Part 2, August 1969,
New York

- [19] Kelly, G.L.:
A Measure of Multiparameter Sensitivity.
Proc. 12th Midwest Circuit Theory Symposium, April 1969,
Austin, Texas

- [20] Schoeffler, J.D.:
The Synthesis of Minimum Sensitivity Networks.
IEEE Trans. CT-11 (1964), S. 271 - 276

- [21] Rosenblum, A.L.; Ghausi, M.S.:
Multiparameter Sensitivity in Active RC Networks.
IEEE Trans. CT-18 (1971), S. 592 - 599

- [22] Shenoi, B.A.:
Optimum Variability Design and Comparative Evaluation
of Thin-Film RC Active Filters.
IEEE Trans. CAS-21 (1974), S. 263 - 267

- [23] Haykin, S.S.; Butler, B.A.:
Multiparameter sensitivity indexes of performance for
linear time-invariant networks.
IEE Proc. 117 (1970), S. 1239 - 1247

- [24] Glesner, M.:
Zur Anwendung nichtlinearer Programmierungsverfahren auf
die Empfindlichkeits- und Toleranzoptimierung elektrischer
Netzwerke.
NTG-Diskussionstagung "Optimierung beim Entwurf von Schal-
tungen", Sept. 1974, Darmstadt

- [25] Manaktala, V.K.; Kelly, G.L.:
Minimization of Sensitivity of Electrical Networks Over
a Frequency Interval.
8th Allerton Conference on Circuit and System Theory,
October 1970, Urbana, Illinois

- [26] Hoffmann, U.; Hofmann, H.:
Einführung in die Optimierung.
Verlag Chemie GmbH, Weinheim/Bergstr. 1971

- [27] Carroll, C.W.:
The Created Response Surface Technique for Optimizing
Nonlinear, Restrained Systems.
Operations Research 9 (1961), S. 169 - 184

- [28] Fiacco, A.V.; McCormick, G.P.:
Nonlinear Programming: Sequential Unconstrained Minimi-
zation Techniques.
John Wiley, New York 1968

- [29] Gaißmaier, B.; Hohnacker, W.; Unbehauen, R.; Wehrhahn, E.:
Optimierungsverfahren für den Schaltungsentwurf, Über-
blick und vergleichende Beurteilung.
Frequenz 29 (1975), S. 112-117 und S. 124 - 133

- [30] Unbehauen, R.:
Synthese elektrischer Netzwerke.
R. Oldenbourg, München 1972

- [31] Glesner, M.; Blum, A.:
Worst-Case Error Analysis of Electrical Networks with
the Aid of Nonlinear Programming Methods.
1974 European Conference on Circuit Theory and Design,
London, July 1974, IEE Conference Publication No. 116
S. 312 - 318

- [32] Orchard, H.J.:
Inductorless Filters.
Electronic Letters 2 (1966), S. 224 und 225

- [33] Klötzner, W.G.:
Statistische Verfahren zur Analyse von Schaltungen.
NTZ 18 (1965), S. 693 - 698

- 34 Reinschke, K.:
Toleranzanalyse elektrischer Schaltungen unter Berücksichtigung statistischer Kompensations- und Überlagerungseffekte.
Nachrichtentechnik 17 (1967), S. 192 - 199
- [35] Becker, D.:
Zur Berechnung von Systemeigenschaften aus den Verteilungsfunktionen der Systembausteine (Methode der Verteilungstransformation).
Frequenz 21 (1967), S. 286 - 292
- [36] v. Blücher, L.:
Zur Toleranzbemessung von Bauelementen in elektronischen Geräten.
Wiss. Ber. AEG-TELEFUNKEN 41 (1968), S. 216 - 222
- [37] Cramér, H.:
The Elements of Probability Theory.
John Wiley, New York 1965
- [38] Huber, B.:
Toleranzanalyse elektrischer Schaltungen zur Ermittlung der Ausfallrate auf statistischer Grundlage.
A.E.Ü. 23 (1969), S. 601 - 604
- [39] Fellingner, J.; Gutsche, H.:
Statistische Toleranz-Analyse elektrischer Netzwerke.
Frequenz 27 (1973), S. 214 - 217
- [40] Pearson, E.S.; Hartley, H.O.:
Biometrika Tables for Statisticians.
Vol. 1, Cambridge University Press, Cambridge 1954
- [41] Elderton, W.P.; Johnson, N.L.:
Systems of Frequency Curves.
Cambridge University Press, Cambridge 1969

- |42| Gutsche, H.:
Ein Beitrag zur Berechnung der Wahrscheinlichkeitsdichten
von Systemfunktionen elektrischer Netzwerke.
Frequenz 27 (1973), S. 45 - 51

- |43| Gröbner, W.; Hofreiter, N.:
Integraltafel.
Zweiter Teil, 4. Auflage
Springer-Verlag, Wien 1966

- |44| Hahn, G.J.; Shapiro, S.S.:
Statistical Models in Engineering.
John Wiley, New York 1967

- |45| Fisz, M.:
Wahrscheinlichkeitsrechnung und Mathematische Statistik.
VEB Deutscher Verlag der Wissenschaften, Berlin 1966

Lebenslauf

- 12.3.1944 geboren in Concepción/Chile als zweites Kind des Landwirts Christian Suiter und seiner Ehefrau Dorothea, geb. Hertlein
- 1950-1961 Schulzeit in Monte-Aguila, Temuco und Valparaíso/Chile. Abschluß mit der Reifeprüfung im Dezember 1961
- 1962-1967 Studium der Elektrotechnik an der "Universidad Técnica Federico Santa María" in Valparaíso. Abschluß mit der Diplomprüfung im Dezember 1967
- 1968-1969 Vertiefungsstudium an der Technischen Universität München als Stipendiat des Deutschen Akademischen Austauschdienstes
- 1969-1970 Projektierungsingenieur bei der Firma Sindelen in Santiago/Chile
- 1971-1972 Projektierungsingenieur bei der Kraftwerk Union in Erlangen
- 1973-1976 Wissenschaftlicher Mitarbeiter der DFG und wissenschaftlicher Assistent am Lehrstuhl für Allgemeine und Theoretische Elektrotechnik der Universität Erlangen-Nürnberg

Copy - Center 2000

852 ERLANGEN · Obere Karlstr. 30
Tel. 09131-26197